S Y N T A C T I C U N I V E R S A L S

A N D S E M A N T I C C O N S T R A I N T S

Statistical and other Comparative Evidence
from the Study of Verbs and Adjectives

by

Martin Hugh Prior

School of Oriental and African Studies

Thesis submitted for the degree of
Doctor of Philosophy
of the University of London

1985

Abstract

This study is concerned with the examination of word-order universals and the attempt to explain the diverse statistical patterns by means of a variety of semantic structures.

Central to my study is a critique (a) of Hawkins's approach, the Universal Consistency Hypothesis, and its concentration upon apparently exceptionless universals, and (b) the Bartsch-Vennemann approach which breaks down structures generally into OPerator-Operand pairs.

In Part I, I introduce a geographically and genetically representative sample of 75 languages, and show that not only are violations to be found to Hawkins's principles of consistency, but that they are statistically predictable. The statistical results are also useful in relating word-order patterns with fusional, agglutinative and isolating morphology: in particular fusion is associated with violations of statistical universals.

In Part II, I concentrate on adjectives and verbs and here the patterns of non-exceptionless universals become important: instead of distinguishing intensional and extensional adjectives after Montague, I identify a semantic principle of Adjunction, which includes some intensional adjectives, over and above those adjectives for which an analysis as (intensional) operators is still appropriate.

The analysis further suggests that the Bartsch-Vennemann identification of operators and operands is often the reverse of what it should be, especially for genitival and verb-object constructions, and I

discuss the implications of this for the verb and its semantic relationship with the subject and object.

Finally I consider determiners, tense and focus as operators upon operators, showing why their syntactic behaviour leads to asymmetric statistical patterns.

Acknowledgements

Many people have been instrumental in the preparation of this study: bearing in mind the number of languages I have either sampled or needed to investigate, acknowledgement of all concerned would surely be a mammoth task.

I would first of all like to express my debt of gratitude to Deirdre Wilson, who first urged me to consider linguistics. Though my decision at the time was undoubtedly a gamble, it was one I have not looked back on.

One of the greatest assets in pursuing research is the stimulation of one's academic colleagues, and I would especially like to thank David Bennett, Kate Burridge, Andrew Carstairs, Annabel Cormack, Malcolm Edwards, David Hatcher, Dick Hayward, Dick Hudson, Mary McIntosh, Baz McKee and Neil Smith.

Over the past few years I have sought - and pestered - a variety of people for their native intuitions and other knowledge of specific languages, and I would like to thank in particular Margaret Bainbridge, Peter Bee, Bruce Biggs, James Bynon, Jack Carnochan, Jeremy Davidson, John Dodgson, Hailu Habtu, David Hawkins, George Hewitt, Martin Jones, Khalid Khattak, David Marshall Lang, Rose Morris, John Okell, Hâle Öztekin, R.H. Robins, Susan Rollin, David Ryecroft, Christopher Shackle, Sho Jiro Sekine, Peter Sherwood, Janig Stephens and John Wells.

I would also like to thank Julian Leslie of the Birkbeck College Department of Statistics for his

assistance in statistical questions. Here I would also like to acknowlege a much longer-standing debt to J. Johnstone and his colleagues in the Department of Econometrics at Manchester University, where I gained my undergraduate degree of BA (Econ). I shall always value greatly the statistical training they provided me.

I would like to express a general thanks to the staff of the School of Oriental and African Studies for their assistance during the course of my research there: in this regard the wide-ranging and unique collection of the SOAS Library has proved indispensable. But in particular I would like to thank Dr Theodora Bynon, my supervisor, who often attempted to struggle through my earlier writings long before they deserved to reach the light of day. Her patience as I put together the present study is especially appreciated.

I must finally thank both my parents for their contribution to the course of my present research: although I have no detailed familiarity with the works of my late father, I was familiar from an early age with his espousal of Polish notation, and the importance he attached to the syntax and ordering of elements in mathematical notation: I was thus no stranger to a consistent VSO language! I am lucky to have in my mother someone who has already successfully completed a PhD, and her understanding has been a tremendous source of strength to me over the past few years.

TABLE OF CONTENTS

Table of Contents (continued)

Table of Contents (continued)

LIST OF TABLES AND PAGES

MAPS

DIAGRAM

Note on Abbreviations

This work attempts to keep in line with the conventions of Greenberg and others in the area of word-order universals. Thus VSO, NG, NA, ND, NNum and NR and the permutations of these letters are used to denote orderings, as illustrated in the text. Note that D, Num and R stand for demonstrative, numeral and relative clause respectively.

Similarly I adhere to phrase-structure conventions in expressions such as NP, VP, PP, etc. Note that PoP stands for postpositional phrase while PP denotes either prepositional phrase or pre/postpositional phrase according to context. CNP stands for 'common noun phrase'.

Other abbreviations, together with some of the above, may be found in the List of Abbreviations.

List of symbols and abbreviations

+	link
Abs	Absolutive (case)
Ag	agent (marking)
all	'all'; item in ordering characteristic, allN/Nall
Asp	aspect
Bene	Benefactive, beneficiary
Cl	classifier
CNP	common noun phrase
Dir	Direct, in particular a case in Hindi
é	ézâfé
EMP	emphasiser
Erg	Ergative (case)
F	feminine
Foc	marker of a focussed item
Gen	Genitive
Imp	Imperative
indet	indeterminate
Instr	Instrumental (case)
Ipf	imperfective (aspectual)
M	masculine
NR	noun followed by a relative clause (ordering characteristic)
Obj	object (marking)
Obl	Oblique, in particular a case in Hindi
Pa	past
Perf	perfect (tense)
Pf	perfective (aspectual)
Pl, pl	plural
PoP	postpositional phrase
PSG	Phrase Structure Grammar
PTQ	Proper Treatment of Quantifiers (Montague)
sm	'small'; item in ordering characteristic smN/Nsm
Top	marker of a topicalised item
-Top	marker of an item that is not topicalised (Tagalog)

NOTE ON GLOSSES

1. Hyphens in the examples indicate morpheme boundaries within words and may or may not co-incide with hyphens in the orthography. Such boundaries will be indicated when it is felt helpful to the exposition:

ev- e (Turkish)
house Dat

evler- e
houses Dat

2. Equals signs (=) indicate that more than one word is used in a gloss to explain a single item:

vîdimus (Latin)
we=have=seen

3. In the following two Persian examples:

dîd- am- esh 'I saw him'
saw lsg him

man Hasân- râ dîd- am
I H. Obj saw lsg
'I saw Hassan'

the features lsg are used for -am, since it is retained in the presence of another item such as man, I. But for -esh the gloss 'him' is used, since it disappears when an item such as an NP is present.

The pronoun may still be used however if this is felt to lead to a clearer explanation, but note that the English pronoun chosen will be appropriate to the context, without any implication for the gender and number system used by the language concerned.

O. Introduction

This study is concerned with the degree to which similar patterns of phrase structure and sentence structure may be observed across independently developing languages, and the degree to which such development is constrained by universal linguistic laws based upon meaning, rather than upon form alone.

We may exemplify the separate requirements of meaning and form by taking syntactically identical pairs of phrases such as 'small elephant' and 'hungry elephant'. For logicians there is a clear difference between the relationship of adjective and noun in each case: in particular the use of the expression 'small' depends critically upon the elephant being described as an elephant, rather than as an animal. There is no such problem with 'hungry'.

This study will consider the extent to which such semantic differences are relevant to natural language. In particular it will look at the degree to which semantic explanation and semantic constraints are relevant to the statistics of word-order universals, and the principal contribution of this work in the area is a statistical survey of 75 languages, sampled with a view to minimising genetic and geographical bias.

In April 1961 a Conference on Languages was held in New York and its proceedings published in book form two years later in 'Language Universals' edited by Joseph Greenberg. Greenberg's own contribution to this book, 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements', marked the start of a series of highly factual and statistical investigations over the past 20 years into the nature of linguistic universals. Greenberg's work was motivated by the general awareness among linguists of the time, that certain consistencies could be observed across languages, in particular that some languages were consistent in placing modifiers before governing items, while others equally consistently placed them after. Greenberg captured these

consistencies by means of a typology based upon the concept of dominant word-order of constructions semantically comparable across languages. Thus Japanese could be typed as SOV/Po/AN: its word-order used mainly subject-object-verb order at clause level, used postpositions rather than prepositions, and placed the semantic equivalent of our adjectives before rather than after the noun. In producing universals such as his fourth universal -- that with 'overwhelmingly greater than chance frequency', languages with normal SOV order are postpositional -- his aim was not so much to provide a theoretical framework to explain such phenomena, as simply to identify what the phenomena were as a basis for further investigations.

It is my hope in this study to contribute towards a theoretical framework which will explain the phenomena observed by Greenberg and others following him.

Now what kinds of explanation may be put forward for the existence of such phenomena across languages? One clear possibility is that of syntactic overlap, in which languages such as Dyirbal treat genitives as a form of adjective, and employ case-agreement with the possessed noun in addition to the genitive marking. It is not surprising then that adjectival position (NA/AN) correlates with genitival position (NG/GN).

Such an account provides essentially for a criss-cross of harmonies, but we might attempt to provide a more abstract explanation, where for example adverbs, adjectives and non-subject NPs are all 'modifiers' and the item linking them with the rest of the utterance the 'head'. We can certainly find languages in which modifiers appear mainly to precede as in Japanese, or to follow as in Thai (SVO) or Welsh (VSO). Perhaps the most important account of this kind is the Natural Serialisation Principle of Bartsch and Vennemann (1972). By treating

such modifiers as Operators and the heads as Operands, they attempt to provide a link with the work of intensional logicians such as Montague, since by treating adjectives as Operators -- among other things -- they can deal with the problems of 'small elephants' where 'small' combines with the meaning or intension of 'elephant' rather than its reference.

But neither account can on its own explain the complexity of the statistical correlations to be found, which I shall exemplify from my own language sample. If we compare the correlations between Pr/Po and NG/GN and between NG/GN and NA/AN we find that the first is not only very high, around 90%, but also pervades the language families. With the latter correlation however, there is still a marked though weaker correlation between language families, but little within. There are then two differences in the statistical patterns, one which is synchronic, the relative strengths of the correlations, and another which could well be diachronic, if the lack of correlation within families reflects a slower change.

We find a situation then in which different explanations are required for different phenomena and my own work will concentrate upon the important point that even if all semantic structures are analysable into operators and their operands, the syntactic items representing such structures need not be so analysed. As will be discussed in chapter 1 (§1.2), this may be seen by comparing

log x, and xy,

in algebraic notation; in the first example, we can undoubtedly treat log, the logarithm of, as an operator and x as the operand. But which is which in the second? The answer is that x and y are both operands

and multiplication or 'times' is the inferred operator. But this does not mean that xy lacks an operator-operand structure, simply because the syntax of algebraic notation permits the omission of the operator in a limited number of expressions.

How is such a distinction relevant to natural language? In the present study I shall focus primarily upon adjectives, and argue that they and adverbs are much more variable than other items in the semantic forms they take, and in some cases something like an xy analysis is appropriate - especially when adjectives are really nouns in apposition. What will emerge is that a concept of adjunct is appropriate which is neither operator nor operand in its relationship to the item it adjoins. This conclusion stems partly from the semantic survey and partly from the statistical analysis carried out in this work. Characteristic of the adjunct relationship are pairs of expressions such as 'strong support' and 'strong supporter', where in each case 'strong' describes 'support', regardless of whether 'support' is the head noun or 'supporter' is. It will be noted that the semantic and syntactic properties are identical in each case: it is only their interaction with the head noun that differs, and I shall argue that the semantic structure is sufficiently vague here to permit the adjunct to refer to the most relevant concept embodied in the head item.

Chapter 1 will discuss the degree to which word-order change is externally and internally motivated, and consider the possibilities for each type, including the analysis of constituent structures. In particular I shall consider alternatives to operator-operand as explanatory devices for internal motivation. Chapters 2 and 3 are concerned with statistical patterns, and chapters 5, 6 and 7 with the semantic patterns. Chapters 8 and 9 attempt to bring the statistical and semantic accounts together. It would be close to the mark to say that the statistical discussions

represent a critique of Hawkins (1979, 1980) and the semantic discussions a critique of Bartsch and Vennemann (1972). The discussions of the statistical patterns additionally lead us to the traditional tripartite typology and its relevance to word-order typology: fusion is associated with inconsistent patterns and isolation with consistent (chapter 4).

Turning to chapters 2 and 3 in more detail, my main argument is strongly in favour of the use of statistical universals: any such universal embodying a statistically significant pattern cannot be over-used since its neglect involves the neglect of something which is there to be explained. In taking this attitude I run counter to the arguments of John Hawkins in 'Implicational Universals as Predictors of Word-Order Change' (Hawkins (1979)): statistical universals are to my mind desirable in themselves, firstly because they indicate that there is something there to explain, and secondly we must note that if the 'exceptionless' universals proposed are not themselves statistically significant - as most of Hawkins's proposals are not - an explanation is readily available, that they represent the interaction of two or more statistically rare events. In this case they will not have the predictive powers claimed for them.

The above remarks do not detract however from the usefulness of Hawkins' statistical analysis, and his Prepositional Noun Modifier Hierarchy provides a valuable insight into what has to be explained. In addition his view that Pr/Po is a more central typological criterion than VO/OV is undoubtedly correct, provided we recognise - as is argued in chapter 4 - that case-marking is even more central when it occurs, giving renewed importance to the traditional tripartite typology.

Noun modifiers will also be my concern throughout chapters 5 to 7 - both in prepositional languages and in postpositional languages - and here my discussion forms a critique of the Bartsch-Vennemann work on

modifiers. I do not wish to criticise their 'Trigger Chain' hypothesis as Hawkins does: I believe that my analysis of variation between and within language families vindicates this to some extent. I am however concerned with their identification of Operator and Operand. Chapter 5 summarises the concept of adjectives as operators, motivated by a dichotomy of intensional versus extensional adjectives, and in chapter 6 use is made of an Adjectival Hierarchy in which types of adjectives - including indefinite and numeral - are distinguished semantically. It emerges that intensional adjectives show a greater tendency to precede in prepositional (Pr) languages than other adjectives, and that to some extent they follow in Po languages. This runs counter to what one expects from the Bartsch-Vennemann approach, where Operators follow in Pr languages and precede in Po languages. Chapter 7 presents the alternative account for adjectives in which a dichotomy of adjunctive versus non-adjuntive is preferred.

Chapters 8 and 9 attempt to draw the statistical and semantic strands together: chapter 8 deals primarily with verbs, showing that if we allow for more than one type of semantic structure, arguments can also be put forward for different types of structure for the verb-object and verb-subject combinations. Thus the verb-object VP is by no means universal. Chapter 9 looks at other phenomena, in particular the behaviour of 'operators-upon-operators' or second-order operators, and shows that their behaviour, in the form of demonstratives and sometimes indefinite determiners, is responsible for a number of important asymmetries in the word-order statistics. The chapter concludes by summarising three statistical patterns observed in this study, and considers how far they can be related to the three semantic patterns examined, that is (a) operator-operand, (b) adjunct-adjoinee, and (c) operators upon operators.

It is my hope in this work, to have provided clear statistical evidence for a semantic framework that can be further developed to relate natural language and logical form.

PART I

SYNTACTIC

UNIVERSALS

1. Problems of Explanation

0.

It has been pointed out by writers such as Greenberg (1963) and Bartsch and Vennemann (1972) that to a large extent syntactic universals are describable in terms of consistent patterns. At the start of section 2 of Greenberg (1963) it is stated:

> Linguists are, in general, familiar with the notion that certain languages tend consistently to put modifying elements before those modified or limited, while others just as consistently do the opposite.

This chapter will look at how far linguists have advanced from this generalisation, and what sort of explanations should be sought.

As mentioned in the Introduction, such observations on the key role of modifiers in ordering patterns have led to two approaches. The first has most notably been pursued by Greenberg in his pioneering work (1963), in which various pairs of constructions displayed harmonies. Such harmonies, which he attributed to syntactic overlap, appeared to be as follows:

(i)	Pr	VS	(verb-subject ordering)
	Po	SV	(subject-verb ordering)
(ii)	Pr	VO	(verb-object)
	Po	OV	(object-verb)
(iii)	Pr	NG	(noun-genitive)
	Po	GN	(genitive-noun)
(iv)	NG	NA	(noun-adjective)
	GN	AN	(adjective-noun)

Here we see under (i) that if a language for its normal or dominant ordering places the verb before the subject (VS or verb-subject), this

will be 'harmonic' with having prepositions (Pr). This occurs because VS languages are overwhelmingly prepositional. On the other hand postpositional (Po) languages are overwhelmingly SV. From such direct harmonies two sets have been postulated: in the first set we have

VS, VO, Pr, NG, NA

which are harmonious with each other, either directly or indirectly, and likewise

SV, OV, Po, GN, AN.

However there is no presumption that any single criterion, such as VO/OV, is central. By contrast Lehmann, Bartsch and Vennemann do make use of VO/OV as a central criterion (see Lehmann (1973), and §6.2 of Bartsch and Vennemann (1972)), while Hawkins (1980) rejects this in favour of Pr/Po. In addition Bartsch and Vennemann relate this centrality of ordering to a more abstract indicator: the ordering of Operators (essentially adjectives, adverbs and non-subject NPs), and their Operands.

The introduction of Operators and Operands is not necessarily in conflict with either the Greenberg or other approaches, since for Vennemann inconsistencies represent a state of syntactic instability. Change will then take the form of a'trigger chain' (Hawkins's term in criticising the concept), so that if a language is changing from XV to VX, this will first trigger a change in the main clause, such as that from OV to VO, and then either subsequently or more slowly, changes in the noun phrase. However as we shall see, Hawkins attacks the concept of a trigger chain, even though he otherwise accepts the Operator-Operand account and prefers instead an alternative mechanism for change, the Universal Consistency Hypothesis, employing unilateral multi-valued implicational universals.

But how do we deal with the situation in which we encounter at least three types of statistical correlations, as exemplified by the following which I shall presently show to be critical:

(i) the correlation between Pr/Po and NG/GN is around 90%, in that some 90% of Pr languages are NG and some 90% of Po languages are GN,

(ii) the correlation between NG/GN and ND/DN is by the same measure around 70%,[1]

(iii) the correlation between ND/DN and NA/AN is asymmetric, in that virtually all ND languages are NA, whilst some 70% of DN languages are AN?

The Natural Serialisation Principle cannot account for such variations on its own, and if we accept Vennemann's approach, the difference must in some way reflect the concomitant mechanisms of the trigger chain. We could then argue that correlations are stronger in main clauses than in the NP, since reactions here to the trigger chain either follow or are slower, though this does not account for the asymmetry. The alternative is that the three statistical patterns reflect three semantic structures perhaps represented by NG/GN (and Pr/Po), ND/DN and NA/AN. In this chapter I shall discuss the two hypotheses, one that such statistical patterns arise from language change triggered by language contact, and the other that such patterns reflect the influence of one or more universal semantic constraints.

1. Internal Factors Versus External

Whether we have syntactic universals in the form of statistical patterns or exceptionless ones, such universals only have a value as a first step towards explanation, but a large part of the effort is bound to be toward determining what is explained.

It is a characteristic of many of the word-order universals that correlations occur to a level of say 80-90% in one if not both directions: thus when we take the correlation between Pr/Po and NG/GN some 80-90% of the Pr languages are NG, whilst a similar proportion of Po languages are GN. On the other hand some 70% of NA languages have noun-demonstrative order (ND), whilst virtually all, if not all, of the AN languages require the preposing of demonstratives (DN). It is not difficult to pick up such patterns with a comparatively small sample of languages, witness the 30 languages investigated by Greenberg (1963).

Having established even among a small sample of languages that quite strong patterns of correlation exist, we may ask what type of explanation will apply? It is here that we may contrast external influences - such as language contact - with internal accounts, such as the identification of modifier and head for many if not all phrasal constructions. Such a correlation is important, since a considerable degree of susceptibility to language contact will in its weakest form represent an element of language-change which an internalised theory will be unable to capture, and in its strongest form will render abstract accounts irrelevant to cross-language study.

On the other hand a theory which discounts the influence of external factors such as language-contact will have to introduce some random element, such as the'Moebius' effect that I consider in my final chapter: a 'Moebius band' is like a belt with a twist in it, so that we

find that prepositionality (Pr) goes with numeral-noun ordering (NumN), NumN ordering goes with adjective-noun ordering (AN), and the latter goes with - postpositionality. It is because of the possibility of such an effect that there is nothing incompatible about a situation in which languages continually seek word-order consistency, yet never reach an equilibrium situation in which all languages adhere either to one word-order pattern or one pattern plus its mirror-image.

It is thus possible that syntactic change can be motivated either internally or externally, but it is of course possible too that language contact and a Moebius effect combine to produce syntactic change, so that the nature of the attempts to restore consistency are induced by language contact. In this case we must turn back to statistical methods, not merely to monitor individual languages but also families of languages: an important part of this present work is the use of a sample of languages chosen as far as possible to be geographically and genetically representative (see next chapter). Certain advantages of such a sample will already be known: a sample which over-represents Indo-European languages will for example over-represent fusional languages. But we must note in addition a very important advantage: the ability to make quite a fine comparison of syntactic variation between and within language families.

Such a comparison has already been illustrated in the Introduction, for correlations with genitival position (NG/GN). A further example is that of VO/OV ordering, which shows the greatest intra-family variation: of the 11 language-groups in my sample represented by two or more languages, nine show variation of VO/OV (Austronesian, all VO, and Central New Guinean, all OV, being the exceptions). Furthermore, as we shall see in my next chapter, such variation correlates in virtually every case with

geographically adjacent families. Such a phenomenon would suggest something like a trigger-chain, as put forward by Vennemann, and could be represented as follows:

$$\underset{(9)}{\text{VO/OV}} \longrightarrow \underset{(7)}{\text{Pr/Po}} \longrightarrow \underset{(6)}{\text{NG/GN}}$$

where the 9, 7 and 6 represent the number of families showing word-order variation for the items concerned. In effect the VO/OV order is sensitive to external influences, and in turn influences the Pr/Po and NG/GN ordering with diminishing effect. However the above ordering has still to be established and the conjecture must therefore be discussed further, as it is in chapter 3.

One thing that is not at all clear, if VO/OV variation is due to external influences, is whether verb-object order in one language directly triggers VO order in a contact language, or whether some other element or elements are pivotal: in Latin, if this can be regarded as an OV language, we also find prepositions, initial question-words and relative clauses following the noun (NR), although these word-order patterns are comparatively rare in OV languages. Maybe it was one or more of these rather than VO/OV that was susceptible to external influences. Whether this is the case or not would require further investigation, but if it were found to be so, it would provide clear evidence that factors other than the phonological erosion of the case system, as suggested in Vennemann (1974 and 1975) and Bean (1976), could have been responsible for an OV-VO shift. The implications of my sample data on the interaction of VO/OV, Pr/Po and object case-marking will be spelt out in chapter 4.

The above discussion of VO/OV ordering, with its recourse to the statistical data available, provides evidence for both internal and external causes of word-order change, and we must therefore

consider more carefully the nature of internal factors. The Natural Serialisation Principle, with its concomitant trigger-chain as a mechanism for change, provides an important starting point for such an examination. Here we may identify for each phrase an Operator-Operand structure, where we find that non-subject NPs, adjectives and adverbs are Operators and their governing items Operands. Such an approach, in which it is the Operand that determines the syntactic status of the phrase as a whole, appears at first sight to provide a very attractive explanation, and correlates VO/OV with NA/AN: according to Bartsch and Vennemann, object NPs act as Operators upon their governing transitive verb since it is the verb not the NP that determines the status of the expression as a verb-phrase (VP). A language such as English with VO/AN ordering is considered to be in a process of change, either to OV or NA.

The above approach might appear to provide a model for the internal factors, when they interact with external factors to direct the course of syntactic change. But does such an interaction adequately account for the statistical data? The main difficulties are those of asymmetric patterns, as exemplified by the correlation between ND/DN and NA/AN, and the Moebius effect described above, and it will immediately become apparent that the statistical data requires a bit of unravelling. The asymmetrical nature of some of the data has long been apparent, most strongly in the correlation between NA/AN and ND/DN, but also in the correlation between the former and NG/GN. Greenberg captured this by treating AN (and VS) as recessive orderings, which did not tolerate inconsistent combinations, such as AN/ND and VS/Po. Hawkins further attempted to account for the asymmetric patterns by means of the (virtually) exceptionless Prepositional Noun Modifier Hierarchy (PrNMH): according to this hierarchy we find, for example, that a Pr/GN language must be AN, and further that a Pr/AN language must be DN. The hierarchy is asymmetric, firstly in that it applies only to Pr languages, and secondly that Pr/NG languages may well be AN, etc (see chapter 3, §3.2).

Now the PrNMH is a very important hypothesis, even though I shall show that the explanations for the placing of items on it are quite heterogeneous: but if we exclude ND/DN and relax the requirement of exceptionlessness, we find a mirror-image postpositional hierarchy (see chapters 6 and 9). We further find that if we add quantifiers to it, such as 'all' and the numerals, it forms part of the 'Moebius' chain that I alluded to in the introduction of this chapter. We therefore require two semantic structures (at least), Operator-Operand and something else, to provide the conflict of influences that leads to the 'Moebius' twist. These two structures will be the subject-matter of my next section. We also need a third structure, embodied most notably in demonstratives and their ordering (ND/DN), to account for the asymmetry in the statistics, and this I shall discuss in section 3 of this chapter.

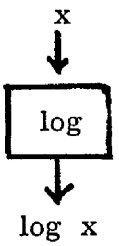
1.2 Semantic Constraints: a first-order level

In the previous section I argued that more than one explanation is required than simply an operator-operand structure if statistics take the form of a 'Moebius' chain, and I shall look at this further in this section. In particular I shall take up the idea I put forward in the Introduction that parallels may be found in natural language to the differences in structure to be found between algebraic expressions such as

$$\log x, \qquad xy.$$

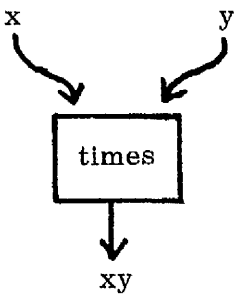
An interesting question also arises, which as far as I know has not been discussed elsewhere, that of what happens with items such as 'very', which are in effect modifiers of modifiers. This represents a second-order level, and in my next section I shall show that these are not necessarily constrained in the same way as first-order modifiers. In particular I shall look at the evidence that such second-order semantic constraints upon word-order are indeed asymmetric. Definite determiners such as demonstratives will also be considered in this group.

So far I have said little about operators, other than mention them in connection with Vennemann's work, and also in the above algebraic example. By way of illustration therefore I would like to look at these algebraic examples a little more closely. In the first case we have the example of a 'log' operator: by placing it before an expression we get a different expression, but an expression which is nevertheless relatable in a regular manner to the first. We thus get



The above diagram expresses the concept of an operator in the form of a box: 'log' has only one input, the x, and one output, log x. Such an expression with only one input, may be termed a monadic operator.

In the expression xy however a different analysis is appropriate:



Here we have an extension of the concept of operator illustrated so far by 'log': instead of having an expression such as log x, relatable by 'log' to the expression x, we have an unexpressed binary operator, relating x, y and xy.

Thus we have the example of two expressions of two elements, whose analysis takes on an entirely different form. In these examples we are able to detect differences on syntactic grounds, in that the 'log' requires a single expression after it, whereas the x and the y require nothing. Yet in the expression xy, the x gets something.

Now to what extent can we find parallels in natural language? I shall first look at Russian, with the following examples:

čto	eto?	what is this?
what	this	
eto	vino	this is wine [2]

The argument for inferring a missing item is straightforward for Russian, since the verb byt', (to) be, is required in other tenses, and in non-finite forms. An alternative argument would be that nominal subject and predicate are identical in form, that is internal structure, and since an expression can scarcely be made up simply of two operators or two operands, something additional must be inferred.

Another form of juxtaposition is that of apposition. The concept of apposition is outlined for English by Quirk et al (1972), and represents the joining of constituents of the same level, where either the reference of the apposites is identical, or the reference of one is contained in the other. We may illustrate this from Hungarian, where the sentence

ez	az	ajtó
this	the	door

means 'this is the door', but in the sentence

ez	az	ajtó	piros
			red

meaning 'this door is red', it has the meaning 'this door'. In each case two NPs are juxtaposed, as we can see from the sentences

ez	piros	'this is red'
az	ajtó piros	'the door is red'.

I have given two examples above, one in which 'X Y' means 'X is Y' and the other in which 'X Y' means in effect 'X which is Y'. These are not the only possibilities, witness the Tiwi language at finite clause level:

purukupaɭli maɭntiŋa	Purukuparli (is) boss
P. boss	
kakiɭituwi jiŋwati	the children (have) honey
children honey	

Here we see that we have to infer either 'is' or 'has' from the context.

The above examples were relatively straightforward, but natural language is not as precise as artificial languages such as algebraic and logical notations need to be: in algebraic notation a minus sign has to be followed by something to negate, but omission of a completing expression in natural language need not lead to an incomplete expression:

John ate his dinner.
John ate.

Thus if we take two Maori expressions:

(he) whare kowhatu	(a) house of stone
(he) whare	(a) house

do we assume that the two common-noun expressions 'whare' and 'kowhatu' are joined by an assumed 'of', or that 'whare' means 'house-of', but in the second case forms a similar 'complete but uncompleted' expression to 'John ate'? I do not propose to give a conclusive answer to such a question in this work, if indeed it is critical. What I do propose is that if we have two types of expression:

$$[\ X \quad [Y]\ ] \quad \text{or} \quad [\ [Y] \quad X\ ]$$

meaning something like 'X of Y', and

$$[\ [X] \quad [Y]\ ]$$

meaning something like 'X which is Y', then the two are subject to different ordering constraints.

I shall term this relationship one of 'semantic apposition', since it captures the semantic content of apposition, even if we find that in semantic terms, X and Y are not constituents of the same level. It must be noted however that this concept is wider than that employed by Quirk et al (1972), where either the reference of the apposites is identical, or the reference of one is contained in the reference of the other.[3] We must rather say that the reference of 'X Y' is contained in the references of both X and Y. Thus 'members (who are) present' precludes members who are not present, and also non-members present, and implies neither that all present are members, nor that all members are present.

Now we come to an important problem of identifying structures:
if we have an expression

$$[\ X \qquad Y\ ]$$

which do we choose, if any, of

$$[\ X \quad [\ Y\ ]\ ],\quad [\ [\ X\ ] \quad Y\ ],\quad [\ [\ X\ ]\ [\ Y\ ]\ ]\ ?$$

We have no problem with

$$[\ [\ \text{members}\ ] \quad [\ \text{present}\ ]\ ]$$

in English, since this is easily relatable to 'members who are present',
nor with

$$[\ \text{pot} \quad [\ \text{inc}\ ]\ ]$$

meaning 'ink-pot' in Welsh ie 'pot of ink', but can we extend this dichotomy
to other expressions ?

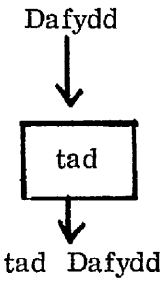
Such examples might be

$$[\ \text{drinking}\ [\ \text{coffee}\ ]\ ] \qquad \text{cf drinking of coffee}$$

$$[\ [\ \text{travelling}\ ]\ [\ \text{teenagers}\ ]\ ] \qquad \text{cf teenagers who are travelling.}$$

but there are plenty of examples such as 'drank coffee', 'of coffee',
'than coffee' and 'that coffee', where it is not clear at all what structure
we are dealing with.

Now there is no difficulty about treating expressions such as 'picture of', 'father of' and 'square root of' as operators: they combine with one expression to form another. It is thus a straightforward matter to interpret the Welsh expression 'tad Dafydd', David's father, as



or alternatively

[tad [Dafydd]]

where in effect the above notation is used to represent

[operator [operand]]

However Bartsch and Vennemann (1972) argue that the reverse should hold, so that we get

[[tad] Dafydd]

representing

[[Operand] Operator]

- where I shall use capital 'O's for their approach. The basis of their argument is that expressions such as genitival NPs are a form of modifier, and should be treated similarly to other modifiers such as

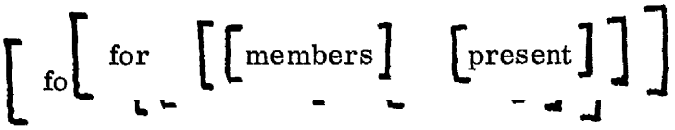
adjectives. Since the treatment of these as operators is an integral part of the work of both Montague and Bartsch and Vennemann, genitive NPs, and indeed most non-subject NPs, are treated as operators.

This approach of Bartsch and Vennemann's, forming the basis of their Natural Serialisation Principle, whereby Operators display a tendency in natural language to consistently precede or follow their Operands, is clearly motivated in part by the assumption that modifier constructions such as adjectival and genitival should be treated similarly. However once I propose in my own approach to identify different semantic structures for adjectives, this motivation is removed. An important feature of the Natural Serialisation Principle is that inconsistencies indicate a state of syntactic change. However I shall show in chapter 2 that not only do adjectives display different diachronic behaviour from genitives, but that such change may well occur towards inconsistency of Operator and Operand ordering, thus running counter to Natural Serialisation.

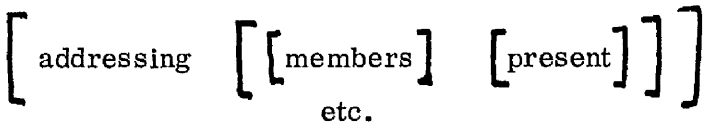
Consequently, since the motivation for like treatment of adjectives and genitives is removed, there should be no objection to my reverting to the treatment of genitives as operands. This will in fact be the approach of this chapter: and the framework I outline will initially make this assumption.

What then can we say about the identification of operator and operand? If they are switched around vis-a-vis the Bartsch-Vennemann approach, then presumably the new operators may take over the role of their Operands in determining the syntactic and perhaps semantic status of the expression as a whole. However there is one thing that may definitely be said about operator-operand structures at a first-order

level: so far we have allowed only for operand expressions to be expanded to contain an 'X which is Y' expression:



This is a very straightforward criterion to apply, and we can easily apply it to a variety of constructions:



An interesting consequence of such a criterion is that pronouns may be treated as operators even where corresponding nominal expressions could only be treated as operands. Thus we find many languages, such as French and Welsh where the possessive 'adjective' precedes while a full nominal possessive expression must follow. This line will not do however in dealing with the English 's construction, and here it would appear that a second-order level is required.

In this section I have focussed upon two types of structure, one in which 'X Y' stands for 'X which is Y' and the other where X is an operator and Y an operand. The identification of the former is far more straightforward and the elements of the structure are commutative in semantic terms. This does not apply to the second structure and we have the additional problem therefore of determining which is operator and which is operand. As a first step I have suggested that if X is restricted in the way it can be expanded, for example a pronominal possessive, it may act as operator even where a corresponding nominal possessor would be treated as an operand. As a result, when we look at languages where it appears that the first constituent of each phrase is the operator, we need only pick out as possibly inconsistent expressions which may be

expanded. Such a criterion however represents neither a necessary nor a sufficient condition for what the operator and what the operand is in any expression, and as will be seen in the discussions of my next section, on a second-order level, it will take the rest of this work to sort out some of the questions raised.

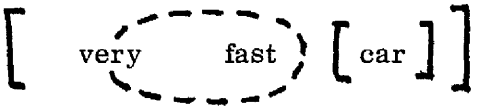
1.3 Semantic Constraints: a second-order level

In this section I wish to look more closely at items which display a more complex structure, rather like modifiers of modifiers, including perhaps adverbs such as 'very', and show that items of such a second-order level have different ordering requirements to items of a first-order level.

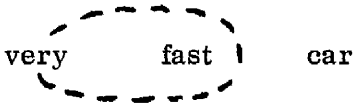
It might reasonably be argued that 'very', represents a modifier of a modifier, and for logicians such as Montague, such items could certainly be considered to be operators upon operators. As we shall see in later chapters, this becomes very apparent from the categorial grammar they employ. Now suppose we treat adjectives as operators upon the noun, as in

[fast [car]]

we must adopt some sort of convention to show the relationship of 'very' to 'fast'. I shall adopt the convention of placing a dashed line around the latter operator:



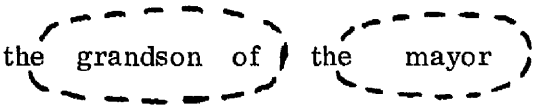
In fact for the purposes of most discussions, it will only be necessary to indicate the region so marked:



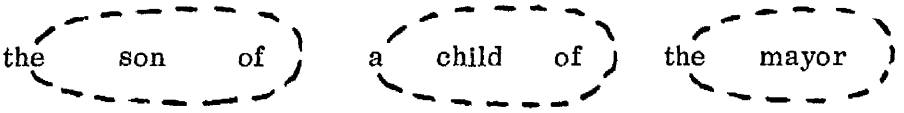
and I shall term this a 'semantic region'.

Although the 'very' is one of the most obvious ad-adjectival adverbs from the point of view of illustration, I shall in fact argue in chapter 7

that 'very' is not such a second-order operator. However there are a number of important areas where such regions are relevant. By way of illustration I shall concentrate here on determiners, particularly definite determiners, such as 'the', 'this' and the pronominal possessives[4]:



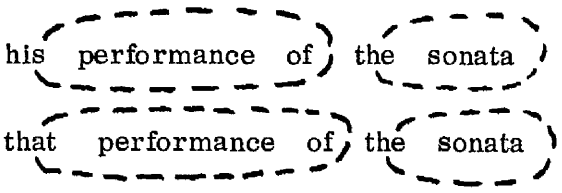
This indicates that everything inside the region is considered as an operator, and that since the region is bound by 'the', a unique result is yielded, unique at least to the discourse. For reasons to be discussed later in this work, it is a matter of indifference whether we put the boundary before or after the 'of'. 'Grandson' does of course mean 'son of a/the child of', but if we spelt this out, we would have to create two regions:



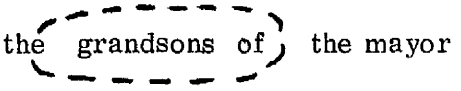
and we have now changed the meaning: the first 'the' now implies that the mayor's child in question has only one son, not that the mayor has only one grandson.

Note that 'mayor' is treated as an operator, even though there is no indication of where the mayor is mayor of. Thus the principle of an operator upon another operator may apply whether the operand is specified or assumed.

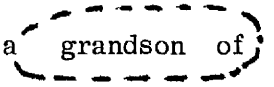
Other determiners may similarly be analysed, though here some additional specification is incorporated:

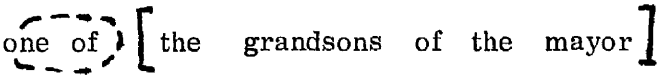


The above examples make use of singular nouns, in contrast to



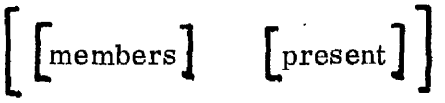
With a plural, clearly 'all of' is implied. Note the contrast in structure implied in the following:





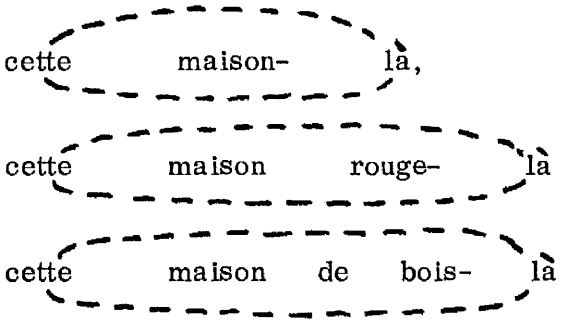
In the first example the indefinite article bounds the region, but in the second a definite article bounds the same region, which now has to be plural, and the 'one of' makes the whole expression indefinite singular. Were there no distinction of number and no definite article, we would be unable to tell whether an indefinite article resembled a definite article in creating a region, or behaved like 'one of', standing separately ~~in~~ with its own region. The same may be said of numerals, and this will be taken up in the last chapter, comparing numeral position (NumN/NNum) and indeed the position of 'all', with other noun 'modifier' orderings.

We may note that the criterion for operators suggested in the last section still applies: in the expression 'the members present', only 'the' can stand as an operator, since 'members present' takes the structure

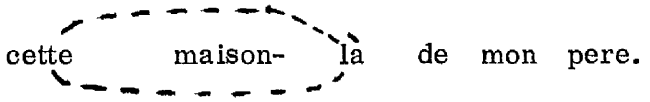


However the data must be looked at further, to determine whether 'the' is a first- or second-order operator.

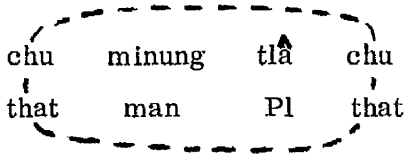
I shall now turn to some French examples, and it would appear that in French and certain other languages, the semantic region is marked at the beginning and at the end:



but



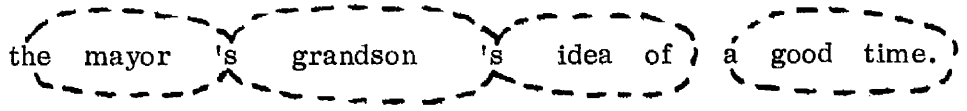
However such languages are not frequent, and indeed the use in French of -ci/là after post-nominal modifiers appears to be avoided if the modifier can be preposed. Such framing constructions nevertheless occur to my knowledge in Malagasy (Austronesian), Songhai (Nilo-Saharan), Kurdish (Indo-European) and certain Sino-Tibetan languages including Bawm:



'those men'

(Bawm is included in my language sample, see next chapter.)

Presumably the 's construction also creates a semantic region, framing it along with the determiner of the possessor NP:



We thus see that there may be various ways in which a semantic region may be framed. In addition, if a head noun always comes first

in the noun phrase, or last in it, this will mark one boundary of the region and the demonstrative or other determiner the other boundary. We may see this in Turkish:

bu	otomobil	terkelek-	ler-	i
this	car	wheel	Pl	its

'these car-wheels'

otomobil-	in	bu	terkelekleri
car	Gen		

'these wheels of the car'

We thus see a strong motivation for putting the head noun and the demonstrative at the periphery of the possessed expression, and the correlation this suggests for ND/DN with both NG/GN and NA/AN will be discussed in the last chapter of this work.

Another area where we might apply such regions is that of tense. In this regard I would like to make three points. Evidence for such an approach firstly occurs when we compare:

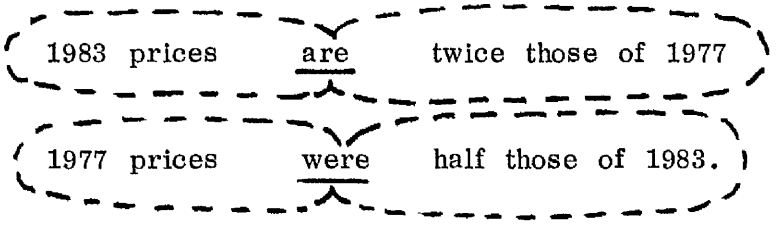
? He thought Sydney was the capital of Australia.

? But I knew that Canberra is the capital.

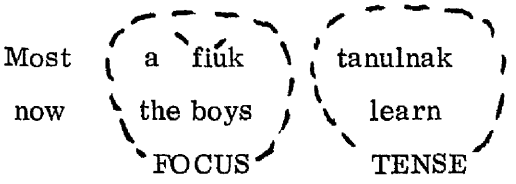
The boundaries suggested, particularly the first, may well be revised on closer investigation, but the point to be made is that even though the tensed verb appears to be head of the whole of the VP expression,

the tense itself may in fact apply to only part of it. Note that similar treatment of tense and demonstratives would accord with Comrie's (1976) proposal that tense is a form of deixis.

Secondly it may be argued that in a language such as English, which appears to split a finite clause into subject NP and VP, the finite tense may create a semantic region for the subject as well as part or all of the predicate:

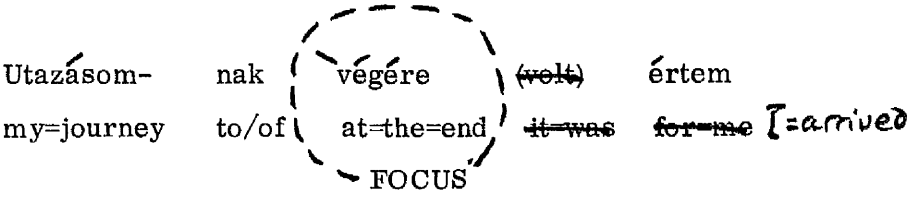


Just as the definiteness or indefiniteness of an NP pertains to a part of the expression, that is the CNP[5], rather than to the whole NP, we find that tense pertains not to the whole of a finite clause, but to the demarcation of subject and predicate. This is the main point to be made at this stage and the precise demarcation of the regions concerned may well require revision on further investigation. Indeed it may well be that only one region is required, especially when we have subject-auxiliary 'inversion'. Circumstantial evidence however for regarding the tensed item as the start of a region comes from languages where it is the focussed item that comes in front of the verb. We may illustrate this with Hungarian (where a high-fall separates topic and focus)[6].



'Now the boys are learning'

I have here treated the focussed expression as a region and evidence for this comes from the fact that only part of a nominal concept need be focussed[7]:



'I am (was) now at the end of my journey'

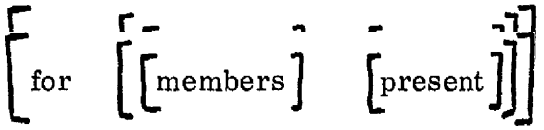
Here we find that the nominal concept 'end of the road' is broken into two nominal sub-concepts: the focussed 'end' and the topicalised 'road'. Now if the possessed expression is to be treated as an operator, focus must be a second-order operator to operate upon it.

The above evidence suggests firstly that if the focussed expression does not co-incide with the verb, it stands outside the tensed region. Secondly we see that the verb must occur early in the tensed region, suggesting that in declarative expressions at least, the subject of an NP-VP clause stands either outside the tensed region, perhaps as a topic, or in a separate implicitly tensed region.

The above discussion of the role of tense and focus has touched upon work that is by no means complete for languages such as Hungarian, or even English, and little in this work will indicate whether it provides a useful addition to the frameworks, such as the Extended Standard Theory, already employed. In my last chapter I shall show however that there appear to be similarities in the ordering constraints on the semantic regions exemplified in this present section. In particular the start of a region is more difficult to infer than its close.

The idea of an 'operator upon an operator' would appear at first sight to be an intimidating concept, but in fact its application appears to be less awesome: diagrammatically it is easily represented as the boundary of a region and it appears to apply equally to a complete expression, or to part of an expression. Thus it is less constrained than a first-order operator, which requires a complete expression for an operand, and may therefore be a rather more difficult concept to identify.

So far I have looked at three structures: operator-operand, semantic apposition and operators-upon-operators, and the question arises of whether we can by analogy deduce 'operators in apposition'. Such structures would presumably capture the semantic content of expressions such as 'buyer and seller of antique furniture' and 'bought the forgery and sold it at a profit', where the operand - in semantic terms at least - has been underlined. It is not my intention to examine in detail the employment of such a device: it would appear to be related both to gapping and anaphora. There is already valuable research in both areas. Their mention here is relevant firstly in introducing them as a natural complement to the set of four semantic structures I have proposed in this work. However by filling in this fourth part of the jigsaw we now have an excessively powerful system, in which there is no firm yardstick by which we can say for any operator-operand structure, which is operator and which is operand. We have already noted the example



in which there is no difficulty about the structure for 'members present', since this clearly means 'members who are present'. At a first-order level, there is no difficulty about making 'for' the operator, since we have not allowed for a complex expression being operator, such as 'members present'. It is only the fourth provision that would permit

this: not only can we have 'members' as an operator in 'for members', but also a construction such as

$$\left[\left[\text{members}\right]\ \left[\text{present}\right]\right]$$

As a result it is indeterminate whether such NPs are operator or operand in such a PP. There is nothing necessarily wrong with this: it may be that though it is all totally arbitrary for any individual construction, once we choose a 'yardstick', everything else will follow.

Further investigation might well suggest that the yardstick be at the highest sentential level: perhaps a tensed semantic region should always be treated as operator upon a focussed expression. In this work however I propose to keep the yardstick more in line with the investigations in the present work: since both Montague and Vennemann agree that adjectives operate upon their noun - intensional adjectives at least (see chapter 5) - I shall adopt this as the yardstick. In effect whatever construction is deemed appropriate for 'intensional' adjectives and their noun, we shall rule out the approach whereby nouns operate upon their intensional adjectives.

2. Statistical requirements: sampling data

0. It is self-evident that since the problems of syntactic universals manifest themselves in a statistical form, part of my analysis must proceed by a statistical route. This chapter and the next are primarily concerned with the methodological problems of this route: the present chapter will be concerned with the sample I have selected of 75 languages and the next with the wider application of this and other samples. Section 1 of this chapter will introduce the sample, discussing the problems of geographical bias and the steps taken to eliminate it. Section 2 will consider briefly the question of statistical accuracy in such a sample. Here I shall focus upon one statistic, that of NA/AN, and discuss the proposition that human languages are predominantly NA, and the difficulties that are in fact to be encountered in attempting to sustain such a hypothesis. Such a discussion is essentially synchronic, in contrast to that of section 2.3, where I shall show how a diachronic element can be added to the discussion, even though the data itself is synchronic - apart that is from knowledge of the relatedness of the languages concerned. In particular here I shall identify evidence that influences upon the ordering of adjectives, NA/AN, are if they exist much more slow-moving in their effect than those of other correlations.

This discussion is an important one, in view of the subject matter of the following chapters, in which I assert the importance of statistical universals even in situations such as those that Hawkins (1979, 1980) points out, where we might appear to be faced with exceptionless universals. In that chapter, and in the last one of this work, such diachronic considerations form an important background to my discussion of the direct and indirect correlations found in statistical universals, and their explanation.

2.1 Allowance for language families

In this section I shall present the languages sampled for this work. The overall characteristics of the sample may be seen in the following table, in which languages are listed in a 'Eurofugal' order, with basic typologies:

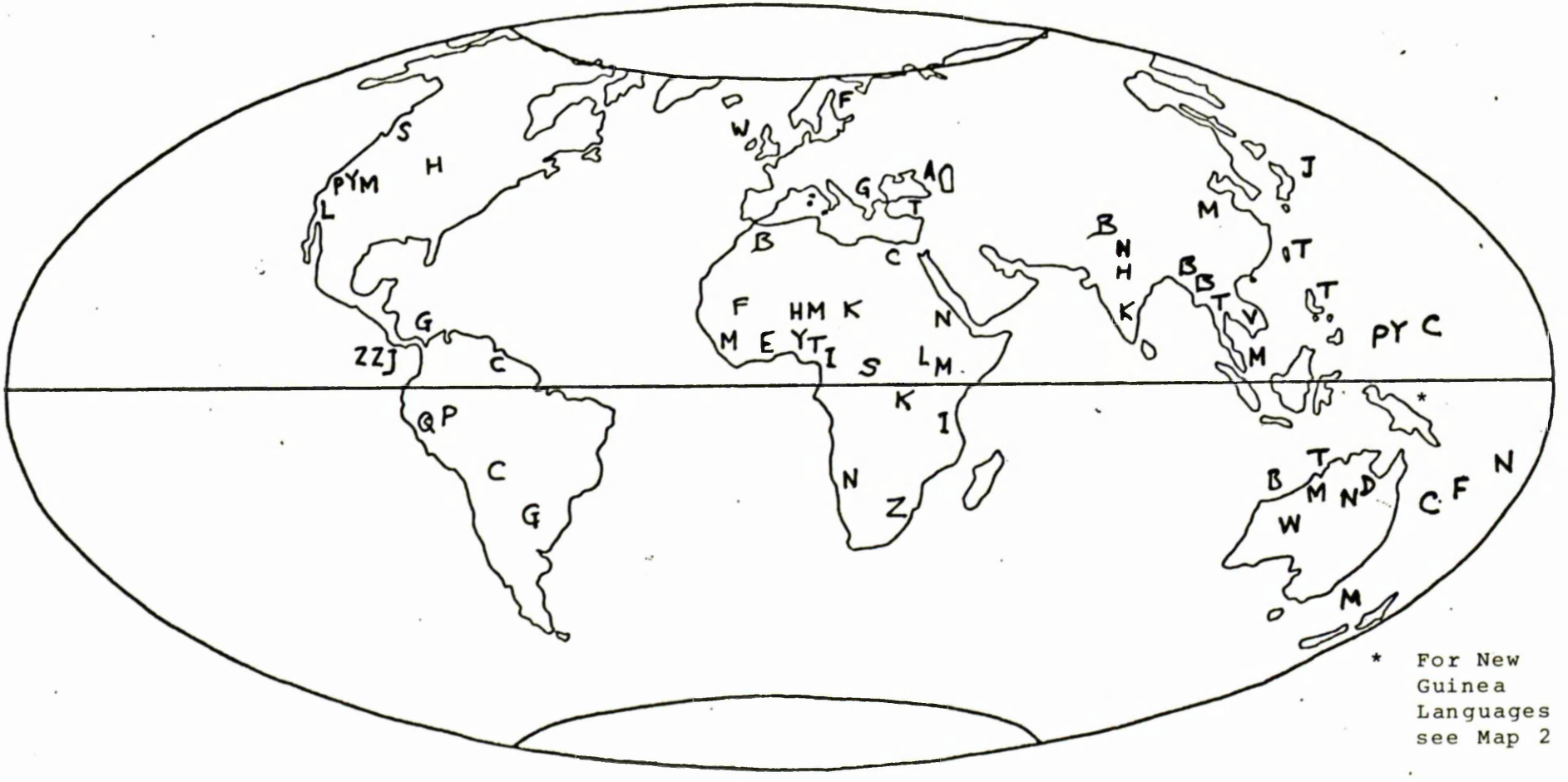
TABLE 2.1

Area	Total	VSO/VOS	SVOPr	SVOPo	SOV/OVS
European	3	1	1	1	-
Asia (excl. Austronesian)	12	-	3	-	9
African	19	2	9	2	6
Australian	6	1	2	-	3
Austronesian	10	7	3	-	-
(Sub-total)	(50)	(11)	(18)	(3)	(18)
Amerindian	15	4	1	3	7
New Guinea (Indo-Pacific)	10	-	2	-	8
Total	75	15	21	6	33

Thus we have listed 15 verb-initial languages, 27 SVO languages and 33 object-verb languages. The breakdown is geographical, except for the extraction of the Austronesian languages. When languages varied their word-order for tense or aspect, the languages concerned were classified in accordance with their order for the Past Perfective form (see Welsh, Ewe and Lugbara below).

The sample was selected with the express objective of avoiding geographical and genetic bias. However, since the Indo-Pacific and Amerindian languages were less accessible than the other groups, less were selected in number than 'should have been': their selection and that of others will be discussed below.

Map 1. Language Sample



* For New Guinea Languages see Map 2

The languages selected were as follows:

EUROPEAN

Indo-European:	Greek (SVO), Welsh (VSO)
Uralic:	Finnish (SVOPo)

ASIAN

Altaic:	Japanese, Turkish (SOV)
Austroasiatic:	Vietnamese (SVO)
Caucasian:	Abkhaz (SOV)
Dravidian:	Kannada (SOV)
Indo-European:	Hindi (SOV)
Kam-Tai:	Thai (SVO)
Sino-Tibetan:	Mandarin (SVO), Burmese, Bawm,Newari (SOV)
other:	Burushaski (SOV)

AFRICAN

Afroasiatic:	Berber (VSO), Cairene (Arabic), Hausa, Margi (SVO), Iraqw (SOVPr)
Khoisan:	Nama (SOV)
Niger-Kordofanian:	Fulani, Kinyarwanda, Sango, Tiv, Yoruba, Zulu (SVO), Ewe (SVOPo), Ijo, Mandingo (SOV)
Nilo-Saharan:	Masai (VSO), Lugbara (SVOPo), Kanuri, Nubian (SOV)

AUSTRALIAN

Pama-Nyungan:	Wati (SOV)
other:	Bardi (?VSOPo), Maung, Tiwi (SVO), Djingili, Ngandi (SOV)

AUSTRONESIAN

E Oceanic	Cemuhi, Fijian (VOS), Maori, Niue (VSO)
other:	Tagalog, Tsou (?VOS), Yapese (VSO), Chamorro, Malay, Palauan (SVO)

AMERINDIAN

Macro-Algonquian:	Yurok (SVOPr)
Andean-Equatorial:	Cayuvava (VOS), Guaraní (SVOPo), Piro (?OVS)
	Quechua (SOV)
Aztec-Tanoan:	Luiseño (SVOPo)
Macro-Chibchan:	Guaymí (SOV)
Ge-Pano-Carib:	Carib (SOV)
Hokan:	Pomo, E (SOV)
Oto-Manguean:	Zapotec (VSO)
Penutian:	Jacaltec (VSO), Zoque (SVOPo), Maidu (SOV)
Salish:	Shuswap (VSO)
Macro-Siouan:	Hidatsa (SOV)

NEW GUINEAN (Indo-Pacific)

Bougainville:	Nasioi (SOV)
Ganglau:	Siroi (SOV)
New Britain:	Sulka (SVO)
Central N. Guinea:	Daga, Fore, Kâte, Marind, Sentani, Waskia (SOV)
Northern N. Guinea:	Olo (SVO)

The above languages are shown on the map above, indicated by their initial letter. Thus for West Africa the letters M, E, Y, T, I and S stand respectively for Mangingo, Ewe, Yoruba, Tiv, Ijo and Sango, while those above them, H, M and K, stand for Hausa, Margi and Kanuri. These languages are also listed in Appendices I-II.

Unless otherwise stated, OV languages are postpositional and VO prepositional. Note that Welsh, Ewe and Lugbara use an auxiliary for imperfective forms, so that if we treat the lexical verb as V, we get in

these instances SVO for Welsh and SOV for the other two. In Greenberg there was some uncertainty as to whether Mandarin was Pr or Po, and Hawkins considered Mandarin to have both SVO and SOV order. My classification of Mandarin as SVOPr will be discussed in section 3.1.

The above sample, with its stratification by language family, used a method of avoiding genetic bias following Bell (1978). Here the sampling unit was a group of languages with an estimated separate existence of at least 3,500 years. Bell then provided his own estimated list of such units per major language grouping. I list these below, together with the number required if approximately one in 4.75 are to be selected:

TABLE 2.2

Linguistic stock	Number of sampling units	One in 4.75
Dravidian	1	0.2
Eurasiatic (Ural-Altaic)	13	2.7
Indo-European	12	2.5
Nilo-Saharan	18	3.8
Niger-Kordofanian	44	9.3
Afroasiatic	23	4.8
Khoisan	5	1.1
Amerind	est. 150	31.6
Na-Dene	4	0.8
Austric	ca. 55	11.6
Indo-Pacific	est. 100	21.1
Australian	ca. 27	5.7
Sino-Tibetan	ca. 20	4.2
Ibero-Caucasian	4	0.8
Ket	1	0.2
Burushaski	1	0.2
TOTAL	478	100.6

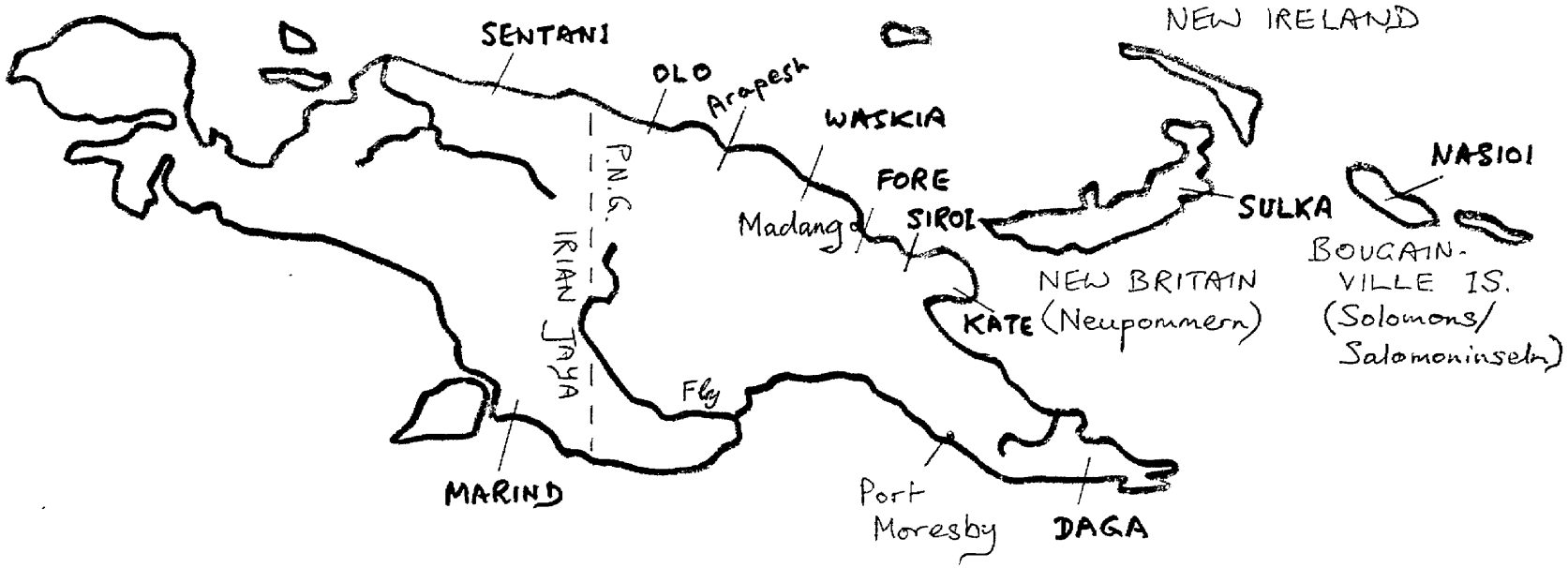
As can be seen, the figure of 4.75 yields a total of approximately 100.

The sample was then stratified in accordance with the one-in-4.75 selection. This was simply done by taking the figure in column 3 of the above table and rounding it to the nearest whole number. Thus Indo-European, with 12 sampling units, was allocated a stratum of three languages, since 12 divided by 4.75 gives 2.56, to which 3 is the nearest whole number. 'Austric', with 55 sampling units, was similarly allocated a stratum of 12 languages, though this was split between Austronesian (10), Austroasiatic (1) and Kam-Tai (1). Dravidian, Ket and Burushaski were too small to allocate a stratum. A pool was therefore made of these, plus Ainu, Basque, Chukchee-Kamchatkan, Gilyak and Yukaghir — other languages which could not be fitted into any of the larger language stocks. Since all of these had the characteristics SOV, Po, GN, AN, except Basque, which had NA, Kannada, a Dravidian language, and Burushaski were selected as the most representative and accessible from this pool of nine units.

In examining the Amerindian, Na-Dene and Indo-Pacific stocks, we find that Amerindian (incl. Na-Dene) and Indo-Pacific are entitled to strata of 32 and 21 respectively. However these languages were simply not accessible in these numbers. Some 10% of both groupings were simply not available, reducing the two strata to around 30 and 20 languages each: it was then decided to sample only 15 and 10 languages for each group, ie only half what should have been.

The final stage was to select languages from within each stratum. This was done in consultation with Voegelin and Voegelin (1977), and selection was then made on the basis of (a) accessibility of the languages concerned, and (b) judgement as to the genetic representativeness of the languages within each stratum. Since Bell gave no indication of which groups of languages had had 3,500 years of independent existence, or the basis of such estimates, this had to be guessed from the arrangement and

Map 2. New Guinea



grouping of the languages in Voegelin and Voegelin. In general, any grouping of languages or individual languages lineally co-ordinate to them was taken as a sampling unit if the total number of such units corresponded roughly to Bell's estimate.

The fact that half the Amerindian and Indo-Pacific languages were sampled that should have been means that these languages may be treated as 2% and others as 1%, since the total then adds up to 100%. We now find that 19% of the sample are VSO/VOS, 24% are SVOPr, 9% are SVOPo and 48% are SOV/OVS. We find in effect that though some 42 of the 75 languages sampled were VO, because of the preponderance of OV languages among Amerindian and Indo-Pacific languages, the proportion reduced to 52% when this weighted percentage was calculated. In this work it has been found that for overall percentages such a weighting is crucial, as in the above example. However when dealing with correlations, such weighting rarely affects the result and makes calculations more complex. The practice in this work then is to type all Indo-Pacific and Amerindian languages in capital letters, when being considered as members of the sample, and to make separate checks for the differently weighted languages. Such separate checks will from time to time appear in the notes.

The above sample could be described as somewhere between a quota sample and a 'stratified judgement sample'. The characteristic of a quota sample is that once the quotas have been reached for each sub-grouping of the sample, that is sufficient. The quotas resemble the strata of more systematic sampling, however the difference lies in the element of judgement: getting the quota is not sufficient, so a judgement must be made about the representativeness, though clearly the ability to make such a judgement depends upon familiarity with the data.

2.2 Synchronic evidence from synchronic data

In the previous section I concluded my outline of the sample of the present work with the comment that the sample was somewhere between a quota sample and a stratified random sample. In order to illustrate the degree of reliability of this sample I shall consider two questions: (a) the apparent predominance of the noun-adjective (NA) order over the opposite adjective-noun (AN) ordering, and (b) the degree to which NA/AN is correlated with other factors.

These questions are interrelated,in that a predominance of NA could in fact result from a predominance of say VO - in fact this is only marginal. However the two questions have both synchronic and diachronic implications: synchronic in that the data at hand is a statement of the state of affairs of present-day languages, and indeed a sample of the possible such states of affairs, and diachronic in that such a state of affairs may be the cumulative effect of non-random change. I shall attempt to consider the first question in terms of the data on adjectival position.

Looking at the bare statistics, we find that of the 75 languages only 25 are AN. These are:

VOS CAYUVAVA

VSO Bardi, SHUSWAP

SVO Chamorro, Finnish, Greek, Hausa, Mandarin, Maung, Tiwi, YUROK, ZOQUE

SOV Burushaski, CARIB, FORE, Hindi, Ijo, Japanese, Kannada, MAIDU, MARIND, Nama, Newari, QUECHUA, Turkish

where I observe the practice of placing Amerindian and Indo-Pacific languages in capital letters, to show that they may be given twice the weighting of other languages in the sample. ~~Six~~ Seven languages, Sange, Tagalog, Tsou (VSO→VOS), Palauan, LUISENO (SVO), Ngandi (SOV) and PIRO (OVS?) appear not to have a dominant order for adjectival placing.

From this data we may calculate a weighted mean. In the following table the 'Old World' languages are those other than Amerindian and Indo-Pacific:

TABLE 2.3

	NA	indet	AN	Total
'Old World'	29	5	16	50
Amerindian	6	2	7	15
Indo-Pacific	8	0	2	10
TOTAL	43	7	25	75
Weighted %	57%	9%	34%	100%

where each 'Old World' language counts as 1% and the others as 2%.

The above figures suggest then that the languages of the present-day world are predominantly NA, with some regional variation: New Guinean languages are overwhelmingly NA, despite the often suggested correlation of OV with AN. Also the Amerindian languages are more evenly split.

We may therefore ask two questions:

(i) Are the world's present-day languages, of which the above sample purports to be representative, in fact 57% NA or thereabouts?

(ii) Does the above predominance represent a propensity towards NA among human languages?

In finding answers to these questions a critical question arises: how independently are the items sampled? If we are simply trying to make an estimate of the actual distribution of present-day languages, this does not matter, as long as we have rigorous sampling procedures. Suppose NA and AN are evenly distributed. Sampling 75 languages for NA/AN is then a bit like tossing 75 coins once each. We would not be at all surprised if the coins fell heads 38 times[1]. Two-thirds of such experiments would yield between 34 and 41 heads. And nine out of ten between 29 and 46 heads. Similarly for tails. But outside these ranges the probabilities diminish rapidly and there is only about one chance in 300 that the coins would fall heads (or tails) 25 times or less. This is with much stricter sampling procedures than the present 'stratified judgement sample' affords. Thus if it turned out that present-day languages were evenly divided between NA and AN one would expect some serious distortion somewhere.

The next section will attempt to answer question (ii).

2.3 Diachronic results from synchronic data

In this section I shall consider how NA/AN correlates with other characteristics, and attempt to do this within a diachronic perspective. The ability to provide a diachronic analysis from what is essentially synchronic data stems from our assumption that languages sampled have had 3,500 years of separate existence, so that the families to which they belong must be at least that old. The discussion of this section is partly to consider problems of analysis and partly to lay down some methodology, so that the discussion of how NA/AN correlates with other factors is essentially only initiated. I shall in fact concentrate upon the correlation between NA/AN and NG/GN.

So far the data on NG/GN has not been discussed in this chapter: it is tabulated, along with NA/AN, in Table 2.4 (overleaf). We may note that

(i) all 36 Pr languages are NG, except for Mandarin, SULKA, Tiwi, YUROK (GN) and OLO (uncertain),

(ii) all 39 Po languages are GN except Djingili, Kanuri (NG) and Ngandi (uncertain).

This gives us 33 NG languages (weighted 37%, since CAYUVAVA, JACALTEC, SHUSWAP and ZAPOTEC each count as 2%), and 40 GN languages (weighted 60%).

TABLE 2.4a

	NG		GN	
Language family	NA	AN	NA	AN
Austronesian (Palauan, Tagalog, Tsou excluded)	Cemuhi Fijian Malay Maori Niue Yapese	Chamorro	-	-
Niger-Kordofanian (Sango excluded)	Fulani Kinyarwanda Tiv Yoruba Zulu	-	Ewe Mandingo	Ijo
Australian (Ngandi excluded)	Djingili	Maung	Wati	Bardi Tiwi
Afroasiatic	Berber Cairene Iraqw Margi	Hausa	-	-
Sino-Tibetan	-	-	Bawm Burmese	Mandarin Newari
Nilo-Saharan	Kanuri Masai	-	Lugbara Nubian	-
Ural-Altaic	-	-	-	Finnish Japanese Turkish
Indo-European	Welsh	Greek	-	Hindi
other	Thai Vietnamese	-	Abkhaz	Burushaski Kannada Nama
TOTALS	(21)	(4)	(8)	(12)

TABLE 2.4b

	NG		GN	
Language family	NA	AN	NA	AN
Andean-Equatorial (PIRO excluded)	-	CAYUVAVA	GUARANI	QUECHUA
Penutian	JACALTEC	-	-	MAIDU ZOQUE
other Amerindian (LUISENO excluded)	ZAPOTEC	SHUSWAP	GUAYMI HIDATSA POMO, East	CARIB YUROK
(TOTALS)	(2)	(2)	(4)	(5)
Central New Guinea	-	-	DAGA SENTANI KATE WASKIA	FORE MARIND
other New Guinea (OLO excluded)	-	-	NASIOI SIROI SULKA	-
(TOTALS)	(-)	(-)	(7)	(2)
GRAND TOTAL	23	6	19	19

Note: In the above table Palauan, Sango, Tsou, Tagalog (NG) and LUISENO and PIRO (GN) were excluded for indeterminate NA/AN, OLO (NA) was excluded for indeterminate NG/GN and Ngandi was excluded for indeterminate NG/GN and NA/AN.

From Table 2.4 we may derive summary totals which clearly show an overall correlation between NA/AN and NG/GN:

TABLE 2.5

	NG		GN	
	NA	AN	NA	AN
'Old World'	21	4	8	12
Amerindian	2	2	4	5
Indo-Pacific	-	-	7	2
TOTAL	23	6	19	19
Weighted %	25%	8%	30%	26%

Note that the overall pattern, of NG languages being strongly NA and GN languages being more evenly balanced is reversed by two regional figures: all the Indo-Pacific figures are GN (except OLO), yet are still overwhelmingly NA, whilst NG American Indian languages are two each. It is therefore advisable to look more closely at variations within families:

TABLE 2.6

	NG		GN		
	NA	AN	NA	AN	TOTAL
Austronesian	6	1	-	-	7
Niger-Kordofanian	5	-	2	1	8
Australian	1	1	1	2	5
Afroasiatic	4	1	-	-	5
Sino-Tibetan	-	-	2	2	4
Nilo-Saharan	2	-	2	-	4
Ural-Altaic	-	-	-	3	3
Indo-European	1	1	-	1	3
other 'Old World'	2	-	1	3	6
Andean-Equatorial	-	1	1	1	3
Penutian	1	-	-	2	3
other Amerindian	1	1	3	2	7
Central New Guinea	-	-	4	2	6
other Indo-Pacific	-	-	3	-	3
TOTAL	23	6	19	19	67

Now the interesting feature of this table is that in a number of these language families, Niger-Kordofanian, Australian, Nilo-Saharan, Indo-European and Andean-Equatorial, five in all, there are both NG and GN patterns, but where there is a preponderance of NA languages, as in Niger-Kordofanian and Nilo-Saharan, this preponderance is found among both NG and GN languages. It is only in Penutian that all GN languages in the sample are AN and all NG are NA. We cannot of course comment upon those language families with only one representative in the sample.

Another way of looking at the pattern is to take the nine families with both NA and AN ordering: Austronesian, Niger-Kordofanian, Australian, Afroasiatic, Sino-Tibetan, Indo-European, Penutian, Andean-Equatorial and Central New Guinea. When we find that AN languages are all or predominantly GN, as with Sino-Tibetan and Central New Guinean, so are the NA languages. In the case of Austronesian and Afroasiatic, where the NA languages are all NG, so are the AN languages. Again only Penutian is the exception to the pattern.

We thus find a situation in which correlation between NA/AN and NG/GN is greater between strata than within them. To consider a situation where the reverse holds - where the correlation within is greater than the correlation between, we must look at the correlation between NG/GN and Pr/Po, as charted in Table 2.7.

TABLE 2.7a

Language family	NG Pr	NG Po	GN Pr	GN Po
Austronesian	Cemuhi	–	–	–
	Chamorro			
	Fijian			
	Malay			
	Maori			
	Niue			
	Palauan			
	Tagalog			
	Tsou			
	Yapese			
Niger-Kordofanian	Fulani	–	–	Ewe
	Kinyarwanda			Ijo
	Sango			Mandingo
	Tiv			
	Yoruba			
	Zulu			
Australian (Ngandi excluded)	Maung	Djingili	Tiwi	Bardi Wati
Afroasiatic	Berber	-	-	-
	Cairene			
	Hausa			
	Iraqw			
	Margi			
Sino- Tibetan	–	–	Mandarin	Bawm Burmese Newari
Nilo-Saharan	Masai	Kanuri	–	Lugbara Nubian
Ural-Altaic	-	-	-	Finnish Japanese Turkish
Indo-European	Welsh Greek	-	-	Hindi

TABLE 2.7b

Language family	NG Pr	NG Po	GN Pr	GN Po
other 'Old World'	Thai Vietnamese	-	-	Abkhaz Burushaski Kannada Nama
TOTALS	(27)	(2)	(2)	(18)
Andean-Equatorial	CAYUVAVA	-	-	GUARANI PIRO QUECHUA
Penutian	JACALTEC	-	-	MAIDU ZOQUE
other Amerindian	ZAPOTEC SHUSWAP	-	YUROK	CARIB GUAYMI HIDATSA POMO, East LUISENO
	(4)	(-)	(1)	(10)
Central New Guinea	-	-	-	DAGA FORE KATE MARIND SENTANI WASKIA
other Indo-Pacific (OLO excluded)	-	-	SULKA	NASIOI SIROI
TOTALS	(-)	(-)	(1)	(8)
GRAND TOTALS	31	2	4	36

Note: in the above table Ngandi and OLO have been excluded for indeterminate NG/GN. They are Po and Pr respectively.

Here we can see that of the six language families where both NG and GN occur, Niger-Kordofanian, Australian, Nilo-Saharan, Indo-European, Andean-Equatorial and Penutian, five display a correlation between NG/GN and Pr/Po.

It is barely necessary to set out a summary table: all the Niger-Kordofanian languages are either Pr/NG or Po/GN, and similarly for Indo-European, Andean-Equatorial and Penutian. The exceptions to this clear correlation are Australian and Nilo-Saharan, but this could be due to insufficient data. Taking the opposite approach, selecting those languages which have both Pr and Po patterns, Niger-Kordofanian, Australian, Sino-Tibetan, Nilo-Saharan, Indo-European, Andean-Equatorial and Penutian, seven in all, we find that the same five as above display a correlation within the family: the odd group, Sino-Tibetan, does however go against the trend in that its prepositional language Mandarin is not NG but GN, so that all the Sino-Tibetan languages are GN, regardless of whether they are Pr or Po.

This concept of correlation between strata and correlation within strata is important. If we have a situation such as that in the Sino-Tibetan languages, where all four languages are GN, and all but one, Mandarin, are postpositional, it is misleading to say that there is a correlation <u>within</u> the sample: it could well be that there are some NG Sino-Tibetan languages which have escaped sampling, of which 25% are prepositional as well: in this case there can be no correlation. Correlation can then only occur if say Sino-Tibetan is added to Afroasiatic, of which all five languages in the sample are Pr/NG. Here too there is no correlation within the strata: the correlation that occurs when Afroasiatic and Sino-Tibetan are taken together is purely cross-stratal.

Why might such a difference occur, between the NG/GN and NA/AN correlation, which occurs only across strata, and the NG/GN and Pr/Po correlation, which occurs both across and within strata? I shall consider three possibilities. The first is that there is basically no correlation between NA/AN and NG/GN, only the chance co-incidence of a number of families where NG/NA (and perhaps GN/AN) is predominant. This is problematic, if we look at the pattern of the single-language strata:

Table 2.3

	NG		GN		Total
	NA	AN	NA	AN	
Other 'Old World'	2	-	1	3	6
Other Amerindian	1	1	3	2	7
Other Indo-Pacific	-	-	3	-	3
TOTAL	3	1	7	5	16

Here the pattern of NA predominating among NG languages with a more even spread among GN languages recurs amongst these 16 single-item strata. If however we are to accept such a hypothesis, it is clear that we must find an entirely different type of explanation for adjectival ordering than for NG/GN and Pr/Po ordering.

A second possibility is that the syntactic change giving rise to the correlation between NG/GN and NA/AN, more particularly the comparative rarity of NG/AN, is due to influences which do not get picked when looking at groups with 'only' 3,500 years of separate existence. In effect such influences are very slow-moving. Here we do not require a separate explanation for adjectival as opposed to genitival or prepositional modifiers: Vennemann's Trigger Chain will do. But the difficulty arises here when we observe that nine strata in the sample had varying NA/AN orders, compared to seven for Pr/Po and six for NG/GN. If the effect

on NA/AN is the slowest moving, surely there should be fewer strata with such variation, not more?

The third possibility is that the influence upon NA/AN is not slow-moving, but that there are a number of conflicting changes going on: in modern French for example there is supposed to be a shift towards AN despite the NG ordering, and this despite the fact that NG/AN is comparatively rare. In this case the cross-stratal correlations between NG/GN and NA/AN are the net though significant result of such cross-movements. If this is the case we should certainly be thinking of possible separate explanations for adjectival ordering.

3. Statistical requirements: implicational patterns

0.

In order to consider the explanation of syntactic universals in terms of semantic constraints it is important first of all to find the appropriate characterisation of the universals: what value are the statistical universals? Is it appropriate to consider each individual syntactic phenomenon, for example the placing of the genitive, against a variety of apparent 'harmonic' constructions, or can we identify a limited number of principal typological characteristics - in Greenberg's case, three, and in Lehmann's case, one? (Greenberg (1963), Lehmann (1973)).

In my last chapter I outlined the sample I am using in this study, and showed how the sub-division of it into strata of language families permitted both synchronic and diachronic patterns to be inferred. As has been indicated in the introduction and chapter 1, my own research involves the comparison of statistical patterns, and in this chapter therefore I shall argue strongly for the effective exploitation of syntactic universals, that is (i) the establishment of norms, and (ii) the concomitant explanation of two things, not only the norms, but also the exceptions. In this regard my concern is not to distinguish as restrictively as possible between attested and unattested constructions, but rather between statistically significant and non-significant patterns.

This approach is best illustrated through a consideration of the use of such universals in the area of (a) identification of relevant typological characteristics ('terminology'), (b) prediction, and (c) explanation. Exceptionless universals, or the identification of unattested patterns, would appear to be necessary for the effective study of (a) and (b), but

I shall show firstly that even here exceptionless universals are of little value unless such universals also yield statistically significant patterns.

This chapter then will show that the exceptionless patterns or principles of consistency put forward in Hawkins (1979 and 1980) are not in fact statistically significant, and that for the purposes of the type of data covered by Hawkins - and of course the present study - distinction of different statistical patterns, rather than attested and unattested constructions, is the most appropriate approach. This involves not merely the distinction between symmetric and asymmetric patterns for two-variable or 'bi-valued' analyses, but in the case of three-variable analyses, the concept of the 'Prediction Chain' is appropriate, where there is no three-way correlation, but rather two of the variables are related to each other only indirectly, via the third. I shall show that this latter approach adequately accounts for Hawkins's principles of consistency without being vulnerable to refutation through the discovery of a single exception. What is lost in terms of the ability to make predictions is more than gained in terms of identifying what needs to be explained.

In this present study I suggest only two exceptionless universals at a concrete level, (i) the absence of ND/AN languages, ie those which place the descriptive adjective before the noun and the demonstrative after, and (ii) the absence of 'true' passives (see below) among SVOPo languages. Section 1 will open the discussion of exceptionless versus statistical universals and exemplify (ii) above. Section 2 will discuss the problems of Hawkins's statistically non-significant exceptionless universals. Sections 3 and 4 will illustrate the principle of conditional independence in the form of Prediction Chains: 3 will examine the data

from Hawkins's sample and 4 will look further at such chains and certain problems in their use, on the basis of my own sample.

In section 5 I shall show how Prediction Chain and stratal analyses may be combined to produce something like a 'Trigger Chain' and here I shall use the resulting statistical data to defend Vennemann and Lehmann - despite other reservations - against Hawkins's criticisms of the concept he terms the Trigger Chain.

3.1 Exceptionless universals

This section discusses the dichotomy between two types of universal: one which states what always happens, and the other which states what happens significantly more often than not, and is hence a candidate for some kind of norm.

Such universals may be illustrated respectively by the following, taken from the 45 universals put forward by Greenberg in his 1963 paper:

> Universal 21. If some or all adverbs follow the adjective they modify, then the language is one in which the qualifying adjective follows the noun and the verb precedes its object as the dominant order.

and

> Universal 2. In languages with prepositions, the genitive almost always follows the governing noun, while in languages with postpositions it almost always precedes.
>
> (my underlining)

The first is an example of an exceptionless implicational universal, while the second is an example of a statistical universal. One way of illustrating such patterns is the use of the tetrachoric table, in effect a table based upon four cells:

TABLE 3.1

	NG	GN	
Pr	31	4	35
Po	2	36	38
	33	40	73

The above table illustrates the relevant data from my own sample, as set out in Table 2.7. All the languages are listed except Ngandi and OLO, where the dominant NG/GN ordering is indeterminate.

The term 'tetrachoric' derives from the four basic entries, here the 31 PrNG languages, the four PrGN languages, Mandarin, SULKA, Tiwi and YUROK, the two PoNG languages, Djingili and Kanuri, and the 36 PoGN languages.

The term is most notably used by Greenberg (1963), but statisticians frequently use the more specific '2 x 2 Contingency Tables' (ie 'two-by-two...'). The fact that no entries are zero makes this a statistical universal, in contrast to the following contingency table, taken from Greenberg (1963) (TABLE 7)

	AN	NA
Adverb-adjective	11	5
Adjective-Adverb	0	8
Adj-Adv and Adv-Adj	0	2

illustrating Universal 21 stated above. This shows that the hypothesis that AN languages never place adverbs after their adjectives is not rejected. Now if the hypothesised implicational universal is not refuted by data outside Greenberg's sample, we clearly have obtained exceptionless universals. However this question is not so obvious when we are faced with a tetrachoric table in which all cells are filled. Returning to our example for Pr/Po and NG/GN, we may look again at the table, with not only actual entries, but also the values, in brackets, that would most likely occur if no correlation existed:

TABLE 3.2

	NG	GN	Total	
Pr	31 (15.8)	4 (19.2)	35	(35.00)
Po	2 (17.2)	36 (20.8)	38	(38.00)
	33 (33.0)	40 (40.0)	73	

Note that just as 35, the total number of Pr languages, is 47.9% of 73, the total number of languages listed, the bracketed figure of 15.8 for PrNG languages, is 47.9% of the 33 NG languages: a situation in which there is no correlation at all between Pr/Po and NG/GN. Clearly we can only approximate to this, since when tabulating languages we are dealing with round numbers, but it will be clear that the actual distribution in the above table is a long way from what we would expect if there is no correlation.

Such contingency tables are commonly found in the social sciences, especially when each cell represents the number of people doing or being such-and-such, and in the above table a Chi-Squared (approximate) test would be appropriate, since all the bracketed figures are greater than 5.0. For tables involving smaller figures, Fisher's exact test is necessary (see Everitt (1977)), and I shall discuss this in section 3.4, when the need for the test starts to arise.

Above is a brief account of the two concepts of universals. The concept of exceptionless universals has understandably a strong attraction, and this attraction is undoubtedly re-inforced by the fact that of the 24 word-order combinations identified by Greenberg, nine appear not to be attested. Much of this chapter will put up arguments however in favour of the statistical universal and its application, especially as a path to the establishment of exceptionless universals at a more abstract level. It is therefore worth noting at this point Hawkins's three arguments against the 'overuse' of statistical universals. These may be summarised as follows:

(i) To the extent that the available data reveals that certain co-occurrence types are totally unattested, we must attempt to distinguish as restrictively as possible between attested and un-attested co-occurrences.

(ii) Implicational statements of a statistical nature, such as Greenberg makes use of are not the appropriate device for capturing regularities of language: in general the statistical universals involve two parameters and the exceptionless three. The latter are thus more revealing.

(iii) It is a regular finding in word-order studies that attested co-occurrences are significantly fewer than the mathematically possible co-occurrences.

Now during the course of this chapter I shall demonstrate that the above arguments are open to quite serious objections, and it may be noted that the two concepts are not mutually exclusive. We can see this if we consider the following example, taken from the SOV languages in my own sample:

TABLE 3.3

	NA	AN	Total
NG	3 (1.6)	0 (1.4)	3
GN	14 (15.4)	14 (12.6)	28
	17	14	31

From this we might deduce, as did Greenberg (1963) and Hawkins (1979), that if a language is SOV with NG order, it will also have NA: witness the zero entry for NG/AN. This is the hypothesis of Greenberg's Universal 5. This would appear at first sight to constitute evidence for an exceptionless universal, but does it constitute evidence for a statistical universal? This is highly questionable, since three NG languages is not a large number anyway. It is perfectly possible that if we had a larger number of SOV/NG languages, say 30, that they would split between NA and AN in much the same way as any other NG languages, ie around

22 NA languages and 8 AN languages (see Table 3.6). To avoid this situation the exceptionless universals should also be statistical universals, ie something to which we can attach the expression 'with more than chance likelihood'.

The above discussion then has shown that the two types of universal do indeed have common features. However they are in other respects rather different phenomena posing rather different problems. Since my own approach will be to make use of statistical universals in surface phenomena to point the way to exceptionless universals at a more abstract level, it is desirable to spend some time looking at these differences. I shall do so by considering three areas, of terminology, prediction and explanation. To this end I shall focus upon a hypothesised exceptionless universal:

> If a language has SVO and postpositional ordering, it will not have a passive construction, where such a construction turns the object of an active verb into the subject of a passive verb.

This hypothesis orginates from an observation made in Heine (1975) about SVOPo languages in Africa, though he talks about their lacking 'true' passives. In my sample there are six SVOPo languages:

> Ewe, Lugbara, Finnish, LUISENO, ZOQUE and GUARANI.

Outside the sample we may of course note related languages such as Estonian (related to Finnish), Songhai (related to Lugbara), and a number of languages in West Africa related to Ewe. We also have !Kung (Khoisan), Rutulian (Caucasian), Tasmanian (Indo-Pacific according to some conjectures) and Kashmiri (Indo-European). According to Greenberg the Algonquian languages are 'probably' SVOPo, though Hawkins is more firm on this point.

This is an interesting hypothesis, if we look at Table 3.4, where data on VO/OV and Pr/Po are given for each stratum. Of the eleven strata in which three or more languages have been selected, all but two, Austronesian and Central New Guinean, have variation for VO/OV. Of these nine, all but two, Afroasiatic and Ural-Altaic, show variation for Pr/Po, and there appears to be correlation in all seven of these strata between Pr/Po and VO/OV. (Though perhaps with Nilo-Saharan, Andean-Equatorial and Penutian more data would be desirable.) These facts suggest that typological change is a very common phenomenon. Furthermore SVOPo languages are known to occur in six of the eleven major strata of the sample, so that SVOPo syntax would appear to be a regular element in such typological change. If it is the case that SVOPo syntax avoids passives, it would appear then that such typological change has important consequences for the verbal system.

TABLE 3.4

	VO		OV	
	Pr	Po	Pr	Po
Austronesian	All 10	-	-	-
Niger-Kordofanian	Fulani Kinyarwanda Sango Tiv Yoruba Zulu	Ewe	-	Ijo Mandingo
Australian	Maung Tiwi	Bardi (VSO)	-	Djingili Ngandi Wati
Afroasiatic	Berber Hausa Cairene Margi	-	Iraqw	-
Sino-Tibetan	Mandarin	-	-	Bawm Burmese Newari
Nilo-Saharan	Masai	Lugbara	-	Kanuri Nubian
Indo-European (NB. Kashmiri)	Welsh Greek	-	-	Hindi
Ural-Altaic	-	Finnish	-	Japanese Turkish
Andean-Equatorial	CAYUVAVA	GUARANI	-	PIRO QUECHUA
Penutian	JACALTEC	ZOQUE	-	MAIDU
Central N G	-	-	-	All 6
other	7 langs.	1 lang (LUISENO)	-	10 langs
TOTAL	35	7	1	32

Because of the strength of the correlation between VO/OV and Pr/Po this hypothesis is of particular interest in that it supplements a very strong statistical universal with an exceptionless universal, and suggests that we are close to some kind of exceptionless universal at an abstract level. Such an account is valuable too in linking word-order with syntactic attributes such as voice. A possible explanation for the phenomenon would be that in a postpositional language, the verb will combine most naturally with what precedes it, compared with a prepositional language where the verb most naturally combines with what follows. In this case the subject-verb combination parallels that of a passive verb plus agent in prepositional languages, a phenomenon which will be further discussed in chapter 8.

The hypothesis is of further interest then, in that it enables us to consider the possibility that when verb and object are constituents of a VP, an exceptionless universal applies, namely that VO goes with Pr and OV with Po.

The explanation of SVOPo constructions as passive in semantic structure must however be further refined, since by the same argument, VSOPr languages should also be passive in semantic structure, since S is closer to the verb than O. But such languages frequently have passives, most notably in the Polynesian languages.

The above universal or constraint does not affect the so-called impersonal passives observed by Comrie (1981a) for Finno-Ugrian languages, as exemplified in the Finnish example:

hänet	vietiin	kouluu-	n
him (Acc)	one=took	school	to

'He was taken to school'

Nor does the constraint appear to affect nominalisations, as in the following example, also Finnish:

tämä	pöytä	on	puusepän	tekemä
this	table	is	carpenter's	making

'This table has been made by the carpenter.'

Here the last word is a verbal noun, described in Whitney (1956) as Infinitive III.

We must also note in LUISENO that there is a large class of verbs which take -i for a transitive form and -ax for a corresponding intransitive. It is not clear whether we may regard this pattern as productive and we find furthermore that in some pairs the intransitive verb does indeed translate as a passive but in others the transitive translates as a causative. Unfortunately the pairs exemplified in Kroeber and Grace (1960) are limited, mostly to examples of derivative morphology, but we still have the pairs [1]:

neči,	pay	nečax,	be paid
qarí'i,	raise	qarí'ax,	rise
šu'li,	put in	šú'lax,	visit (ie enter)

We may note however that we cannot obtain the straight passive 'be raised' and it would thus appear that if nečax translates as the passive of neči, pay, this stems from the meaning of 'pay' rather than that of the suffix -ax.

If the above hypothesis can be sustained, so that an SVOPo language with passives is linguistically impossible, or at least never motivated, this has important consequences for typological studies. I shall first consider the question of terminology, or the definition of typological characteristics.

It is clear that for an exceptionless universal to be meaningful, we must be precise in defining the terms used in stating the universal concerned. What do we mean by 'dominant', by SVO, and by postpositions? What characterises postpositions? Are they a universal syntactic category? What is a syntactic category? And so on. An important question arises of whether the governing criterion is a syntactic category of pre/postpositions, the construction involving one 'model' item such as instrumental 'with' (or even agentive 'by'), or some rather more abstract concept such as operators and operands.

On the basis of a dozen or so observations (perhaps more in Heine's African language sample, though the usefulness of this may depend upon how closely the languages concerned are related), it will only be possible to get an incomplete answer to these questions. The problem then arises that while there are clear advantages with an exceptionless universal in that it tells us something about what definitions are relevant to syntactic universals, we must have some kind of circularity. I shall therefore give two examples, each from languages with passive-like constructions, one where the Pr/Po status is uncertain (Mandarin), and the other where the uncertain status is that of VO/OV (Kashmiri).

In Mandarin Chinese we find particles which follow NPs with the semantic content of English prepositions[2]:

dì-	shang	'on the floor'
floor	on	

However the resulting 'postpositional' phrases have the same distribution as an NP, as sentence topic, or object of a verb:

Qiáng-	shang	guà-	zhe	yí-	fu	huà
wall	on	hang	Asp	one	Cl	painting

'On the wall is hanging a painting'

Zhāngsān	zài	dì-	shang	shùi
(name)	be=at	floor	on	sleep

'Z-S is sleeping on the floor'

We must therefore note that the expression 'dì-shang' does not directly modify shùi, sleep, but combines with the zài, be at, to form a phrase which goes with shùi. Is this a prepositional phrase modifying the shùi, or a verb phrase in apposition to it? My view is the latter, but this is not necessarily relevant if comparative typology involves semantic criteria, as argued in Greenberg (1963). In effect, by classifying Mandarin as Pr, I am treating it as prepositional without it having any prepositions! This is because the item which combines with a nominal expression to form an expression - be it VP or PP - which describes for example a locative expression, precedes that nominal expression. Inasfar as semantic explanations are being sought for syntactic patterns this seems to be the right approach, but the test must lie in what one can do with such an approach.

Another language whose status as prepositional or postpositional has to be examined carefully is YUROK. Undoubtedly prepositions exist in the language:

yo'	'o'lowo'm	wenepuy	ho	ku	pegʷ‿k
she	gave	her=salmon	to	the	man

'She gave the man her salmon'

On the other hand we find the use of a locative suffix, corresponding to an English preposition:

tepo	-	+	'in a tree'
forest		Loc	

However it is not clear whether such suffixes can be added to any NP, or simply to a noun possibly possessed (R H Robins, pc). We find on the other hand that many English prepositional expressions are represented by what look like adverbial genitives:

ri' k'ew	we-	nesk^{w}i	'near the sand-bar'
sand-bar	its	near	

All the above examples, unlike Mandarin, are undoubtedly adverbial expressions. However YUROK has passives.

I shall now turn to Kashmiri, which is SVOPo, certainly for synthetic tenses:

tami	trowü	sörüy	panünü	naukar-	as	-peth
he(Erg)	left	entirely	his=own	servant	Dat	upon

'He left (to) his servant all his business'.
(Grierson (1911), Ex 1051)

But passives also occur, or at least the following construction is cited by Grierson as a passive:

yih	kitāb	yiyi	takān	chhāpana
this	book	come	quickly	from=printing

'This book will shortly be printed' (Ex 1392).

Such constructions are quite distinct from the ergative construction - even though the intransitive subject and (transitive) object take on the same case in each example - since the passive construction employs an auxiliary yinu, (to) come, plus the ablative of the verbal noun. The difficulty is that in periphrastic constructions we find both VO and OV constructions, if we treat the lexical verb as V. Thus we have:

bŏh	chus-na	chyŏnu	matlab	bōzān
I	am not	your	meaning	understanding

'I do not understand your meaning' (Ex 1837)

In both the Kashmiri and Mandarin examples then, we encounter problems of definition, if we are to sustain the original hypothesis, in the one case as to what constitute Po,

and in the other as to what constitutes SVO. In the case of Mandarin at least, there is nothing wrong with treating the language as Pr, providing we stick to our criteria, and show that the typology produces useful results.

So far so good, but two problems must be raised:

(i) There may be two criteria, one treating Mandarin as Pr and one treating it as Po. It may well require a lot of data, either from a lot of languages, or from variation in a more limited number of languages, to determine which is the more appropriate in any situation.

(ii) We may well still go ahead with the analysis of statistical universals, with the above provisos, but we have still not solved the problem of circularity with the SVOPo hypothesis.

I shall now look at the question of prediction. Here we are coming into the area of Hawkins' work (1979, 1980) where implicational universals are used as predictors. Not merely should we be able to make synchronic comparisons of languages, such as the above, but if the SVOPo hypothesis is sustained, we may also state:

(i) If a family of languages clearly had a postpositional proto-language, and also appeared to have a passive construction, this will point to OV ordering for this proto-language, and the loss of such a passive construction before it could become VO.

(ii) If a language family such as Indo-European uses a wide range of passive constructions, then the possibility cannot be ruled out that the Proto-language was SVOPo.

(iii) If an SVOPr language has a passive construction which it shares with an earlier SOVPo stage, the evidence is that the language went through either an indeterminate stage, or an SOVPr stage, rather than an intermediate SVOPo stage.

It will thus be seen that we not only have a device in such implicational universals of spotlighting cross-language syntactic regularity, but also we have <u>implicational universals as predictors of word-order change,</u> precisely the subject matter and indeed the title of Hawkins' 1979 paper, where he stresses the advantages of exceptionless rather than statistical universals. Both types of universal figure in the first two of his four hypotheses:

> <u>The Universal Consistency Hypothesis (UCH)</u> At each stage in their historical evolution, languages remain consistent with synchronic universal implications.

At first sight this would appear to be tautologous, and the hypothesis with which he contrasts it a logical impossibility:

> The Universal Violation (or Trigger-Chain)[3] Hypothesis (UVH)
>
> Languages violate synchronic implicational universals of the form 'If P, then Q' by evolving * P and ~Q co-occurrences. Such co-occurrences trigger a chain of subsequent word-order changes which re-introduce consistency, as languages acquire whatever Q properties are observed to co-occur with P on syntactic evidence. (where '~Q' stands for 'not Q')

At first sight, this hypothesis seems difficult to test, since we apparently must find a language which 'does what does not happen' and he in fact discovers that Arapesh'does what doesn't happen', but as we shall see, effectively disregards it. However the force of this UVH hypothesis is that when languages observe patterns that are comparatively rare, these patterns are short-lived, and that changes resulting display a trigger-chain pattern. As will be shown in my next section, Hawkins' evidence towards this question is inconclusive, but in fact in my alternative to Hawkins' approach, evidence does appear to exist for something like a trigger-chain.

It may be felt that the main advantage of exceptionless universals is that they constitute explanations. However, if they take the form of zero entries to tetrachoric tables, they themselves must require explanation. It may well be that the absence of passives in SVOPo languages explains certain patterns of syntactic change, however the phenomenon itself requires explanation, and this may well be a difficult task, if all that strikes us is the hard facts of co-occurrence and non-co-occurrence.

Consequently it may well be essential to look at statistical correlations in related areas: this may take the form of attempts to look for languages which are only partly postpositional, and find which postpositions are critical. It would seem from the above evidence that locatives are not.

An important thing about having a single exception is that unless we can immediately explain it, we immediately lose the ability to pin down the definitions for the relevant terminology. We may exemplify this with Arapesh, the one exception available to Hawkins in his hypothesis that Pr/GN languages must be AN. But one of the problems is that Arapesh appears to have only one 'preposition' that can be used with NPs: um, whose principal function is to introduce a sentence-final expression:

a	bǫ́ra	bagabys	um	ulypat
they	cut	floorplanks		house

'they cut up floorplanks for a house'

For more complex prepositional concepts, an anticipatory pronoun must be used:

na	tem	sihah	ig	um	bidubig
3pl	perch	top=side	it		canoe

'they perched on the top of a/the canoe'

Given this highly exceptional pattern for Arapesh, the prepositional status of the language is problematic, and it is only when we come across Sulka, some five hundred miles to the east in New Britain, that we find that this argument cannot hold, since Sulka has plenty of prepositions.

It is now appropriate to look afresh at statistical universals. I shall again look at the three questions of terminology, prediction and explanation. And here I shall consider two questions: (i) what would be the consequences if an SVOPo language were found in which passives existed? (ii) what are the consequences for Hawkins when the language Arapesh on his own admission violates his hypothesis that prepositional GN languages cannot be NA?

Looking at the first question, we find that if all but one SVOPo language lacked passives, there would in fact be problems: because we appear to have a statistically significant phenomenon, we have established a norm. The question must then be asked, what characteristics does the language have which make it exceptional? There may in fact be several possible characteristics, and we have no means of identifying which. Furthermore, since the converse does not hold, that all non-SVOPo languages have passives, it may be that the characteristic is shared with other SVOPo languages but in these other languages, passives simply happen not to occur. Consequently, while such languages do not refute the statistical universal, we still have a phenomenon requiring explanation.

As far as prediction is concerned, we find that our exceptions also make predictions impossible, unless we undertake the problematic task of attaching some probability to our predictions (see §3.3). To get around this we must either find an explanation for the 'norm', to see whether such an explanation allows for the exceptions of its own accord, or alternatively try and assess the degree of stability of the structure. We do not know, for example, whether Arapesh is moving towards the position of Sulka, with a number of prepositions, which would suggest that the Pr/GN/NA construction is stable, or away from it - perhaps towards postpositionality, which would suggest instability.

This brings us back to the question of explanation: it is clear from the above discussion that while in some ways an explained statistical universal is as good as an unexplained exceptionless universal, eg for prediction, both are in fact equally in need of explanation, and that the apparent inadequacies of statistical universals should not deter one from seeking their explanation. It is in fact important to stress the similarity of exceptionless and statistical universals in the area of explanation: it is thus probably wise to attempt to find explanations for the absence of SVOPo languages with passives, as if the hypothesis were a very strong statistical universal for which no exception has yet been found. This is of course an extremely difficult task. However the example of Hawkins' hypothesis must be borne in mind: the existence of Arapesh and Sulka appear indeed to render much of his arguments pointless and indeed seem to give weight to his UVH - despite its problematic wording - the opposite of his original intention.

At the outset of this section I stated that the arguments Hawkins put forward for exceptionless universals were open to serious objection and that this would emerge from the discussions of this chapter. Do we then need exceptionless universals at all? So far I have compared and illustrated statistical and possible exceptionless universals in this section, and this with particular reference to terminology, prediction and explanation. It would appear at first sight that considerations of terminology and prediction do indeed require us to distinguish exceptionless and non-exceptionless universals as restrictively as possible, as Hawkins argues. However this does not apply to the question of explanation, since the need to identify exceptionless universals does not excuse us from seeking explanations for statistical universals. Indeed, exceptionless universals are of little value, for either terminology, prediction or explanation, if they are not statistical

universals as well. Consequently it is necessary now to look more closely at statistical universals and their use, if we are at a later stage to make effective use of exceptionless universals, in as far as they can be identified.

3.2 Statistical Universals

In this section I wish to look more closely at statistical universals, and to consider how they are to be most effectively exploited. Hawkins questions the relevance of statistical universals to syntactic change, preferring

(i) exceptionless universals,

(ii) multi-valued implications,

(iii) unilateral implications.

The example of our statistical universal for Pr/Po and NG/GN violates all of these:

TABLE 3.1 (extract)

	NG	GN
Pr	31	4
Po	2	36

using the data from the last section.

On the first count, there are exceptions to the pattern - PrGN and PoNG. On the second, it is bi-valued, in that only Pr/Po and NG/GN are involved. Thirdly it is bilateral: not only are we saying that Pr languages are almost always NG: we are also saying that NG languages are almost always Pr.

An example of such a universal conforming to Hawkins' ideal is the following:

> if a language has postpositional ordering, and the adjective precedes the noun, then the genitive precedes the noun,
> ie Po $\supset$ (AN $\supset$ GN)[4].

TABLE 3.5a

Prepositional Languages

	NA	AN
NG	Berber	Chamorro
	Cairene	Fulani
	Cemuhi	Greek
	Fijian	Hausa
	Iraqw (SOV)	Maung
	Kinyarwanda	CAYUVAVA
	Malay	SHUSWAP
	Maori	(7)
	Margi	
	Masai	
	Niue	
	Thai	
	Tiv	
	Vietnamese	
	Welsh	
	Yapese	
	Yoruba	
	Zulu	
	JACALTEC	
	ZAPOTEC	
	(20)	
GN	SULKA	Mandarin
	(1)	Tiwi
		YUROK
		(3)

Excludes: OLO (Pr/NG~GN/NA), Palauan, Sango, Tagalog, Tsou (Pr/NG/NA~AN).

(Total Pr languages 36)

TABLE 3.5b

Postpositional Languages

	NA	AN
NG	Djingili	-
	Kanuri	
	(2)	
GN	Abkhaz	Bardi (VSO)
	Bawm	Burushaski
	Burmese	Finnish
	Ewe	Hindi
	Lugbara	Ijo
	Mandingo	Japanese
	Nubian	Kannada
	Wati	Nama
	DAGA	Newari
	GUARANI	Turkish
	GUAYMI	CARIB
	HIDATSA	FORE
	KATE	MAIDU
	NASIOI	MARIND
	POMO (E)	QUECHUA
	SENTANI	SIROI
	WASKIA	ZOQUE
	(17)	(17)

Excluded: Ngandi (Po/NG~GN/NA~AN), LUISENO, PIRO (Po/GN/NA~AN).

(Total Po languages 39)

The last part makes use of a kind of logical notation:

$X \supset Y$ 'if a language has the property X it also has the property Y'

ie there are no languages having X but not Y.

The above illustrates a 'bi-valued implication'. A multi-valued implication would be:

$X \supset (Y \supset Z)$ 'if X then if Y then Z'

which may alternatively be expressed as 'if X and Y then Z' since the statement will be false if we find any example of a language with X and Y but not Z. It is not clear how bilaterality applies to multivalued implications, but presumably we would have it with

$Z \supset (Y \supset X)$.

As an example, the implication

$Po \supset (AN \supset GN)$

is true for Hawkins' sample, but the reverse pattern

$GN \supset (AN \supset Po)$

is not, taking the example of Swedish and Danish.

Once a table becomes multi-valued, a tetrachoric table is no longer possible, and the following eight-celled table (octochoric?) illustrates the above:

TABLE 3.6

	Pr		Po		Totals
	NA	AN	NA	AN	
NG	20	7	2	0	29
GN	1	3	17	17	38
Totals	21	10	19	17	67

It may also be termed a 2 x 2 x 2 contingency table. The data comes from my own sample and is detailed in Table 3.5.

Here we immediately see the basis of Hawkins's hypothesis, since there is a zero entry for Po/NG/AN. We may also observe, looking at the Pr side, Hawkins's statistical universal:

> If a language has Pr order, and if the adjective follows the noun, the genitive will (with overwhelming probability) follow the noun.

The above is statistical because there are in fact languages with Pr/GN/NA ordering: SULKA in my sample and Arapesh in Hawkins's. I shall express such a qualified implication as follows:

$$Pr \supset (NA \dot{\supset} NG)$$

The difficulty however with the above two universals is that it is not at all clear from the above data that they have any statistical value: of the two Po/NG languages in the above table, both are NA. This compares with the 27 Pr/NG languages, of which 20 are NA and 7 AN (excluding Palauan, Tagalog and Tsou, for which no dominant ordering, either

NA or AN, could be established). It is not at all clear from the above data that the absence of Po/NG/AN languages is not simply due to a situation in which two-thirds of NG languages are NA, and two of them have been picked out in the sample.

To understand why Hawkins came forward with such universals, of which we may list the first four:

I. SOV ⊃ (AN ⊃ GN)
II. VSO ⊃ (NA ⊃ NG)
III. Pr ⊃ (NA $\dot{\supset}$ NG), exception Arapesh
IV. Po ⊃ (AN ⊃ GN),

where VSO has the same influence as Pr and SOV as Po, we must look at the figures Hawkins was using (see Hawkins (1979)).

The data is as follows, with language groups underlined, and the figures after each type showing languages + language groups:

1.	VSO/Pr/NG/NA. (15 + 4)	Celtic languages; Hebrew, Aramaic, Arabic, Ancient Egyptian, Berber; Nandi, Masai, Lotuko, Turkana, Didinga; Polynesian languages and other Austronesian languages; Chinook, Tsimshian; Zapotec, Chinantec, Mixtec, and other Oto-Manguean languages.
2.	VSO/Pr/NG/AN. (4 + 1)	Tagabili and other Phillipine Austronesian languages; Kwakiutl, Quileute, Xinca.
3.	VSO/Pr/GN/AN. (1 + 0)	Milpa Alta Nahuatl.
4.	VSO/Pr/GN/NA.	none
5.	VSO/Po/NG/NA.	none
6.	VSO/Po/NG/AN.	none
7.	VSO/Po/GN/AN. (1 + 0)	Papago
8.	VSO/Po/GN/NA.	none
9.	SVO/Pr/NG/NA.	Romance languages, Albanian; West Atlantic languages, Yoruba, Edo group, most languages of the Benue Congo group including almost all Bantu languages; Shilluk, Acholi, Bari, most languages of the Chad group of Hamito-Semitic exc. Hausa; Neo-Syriac, Khasi, Nicobarese, Khmer, Vietnamese, all Thai languages exc. Khamti; many Austronesian languages including Malay; Subtiaba.
10.	SVO/Pr/NG/AN. (7 + 1)	Modern Greek, Dutch, Icelandic, Slavonic, Efik, Kredj, Maya, Papiamento.
11.	SVO/Pr/GN/AN. (2 + 0)	Swedish, Danish.
12.	SVO/Pr/GN/NA. (1 + 0)	Arapesh (New Guinea).

13.	SVO/Po/NG/NA.	none
14.	SVO/Po/NG/AN.	none
15.	SVO/Po/GN/AN. (4 + 2)	Finnish, Estonian, Algonquian, Zoque, Rutulian and other Daghestan languages in the Caucasus.
16.	SVO/Po/GN/NA. (8 + 1)	Most Voltaic languages, Kru, Twi, Gã, Guang, Ewe, Songhai, Tonkawa, Guarani.
17.	SOV/Pr/NG/NA. (4 + 0)	Persian, Iraqw (Cushitic), Khamti (Thai), Akkadian.
18.	SOV/Pr/NG/AN.	none
19.	SOV/Pr/GN/AN. (1 + 0)	Amharic
20.	SOV/Pr/GN/NA.	none
21.	SOV/Po/NG/NA. (6 + 1)	Sumerian, Elamite, 'Galla', Kanuri, Teda, Kamilaroi and other southeastern Australian languages.
22.	SOV/Po/NG/AN.	none
23.	SOV/Po/GN/AN. (22 + 7)	Hindi, Bengali, and other Aryan languages of India, Modern Armenian, Finno-Ugric exc. Finnish group, Altaic, Yukaghir, Palaeo-Siberian, Korean, Ainu, Japanese, Gafat, Harari, Sidamo, Chamir, Bedauye, Nama Hottentot; Khinalug, Abkhaz and other Caucasian languages; Burushaski, Dravidian; Newari and other Sino-Tibetan languages; Marind-Anim (New Guinea), Navajo, Maidu, Quechua; Ijo.
24.	SOV/Po/GN/NA. (22 + 2)	Basque, Hurrian, Urartian, Nubian, Kunama, Fur, Sandawe, Lushai, Classical Tibetan, Makasai, Bunak (Timor), Kâte (New Guinea), most Australian languages, Haida, Tlingit, Zuni, Chitimacha, Tunica, Lenca, Matagalpa, Cuna, Chibcha, Warrau, most Mande languages.

If we look at the appropriate tetrachoric tables for each of his hypothesised universals we get the following pattern, with Hawkins' sample shown first and then my own:

TABLE 3.7

			Hawkins' sample		Own sample	
			NA	AN	NA	AN
I.	SOV⊃(AN⊃GN)	SOV/NG	11	0!	3	0!
		SOV/GN	24	30	14	14
II.	VSO⊃(NA⊃NG)	VSO/NG	19	5	8	1
		VSO/GN	0!	2	0!	1
III.	Pr⊃(NA$\dot{\supset}$NG)	PrNG	42	13	20	7
		PrGN	1!	4	1!	3
IV.	Po⊃(AN⊃GN)	PoNG	7	0!	2	0!
		PoGN	33	36	17	17

Cells marked with a ! are those precluded or treated as very rare by Hawkins. Data from my own sample is detailed in Tables 3.33, 3.34, 3.5a and 3.5b, and even though the figures from my own sample show little indication that Hawkins' hypotheses are worth stating, Hawkins' own sample suggests that the attempt is worth making, at least for SOV and Po languages: we can see very clearly that PoNG and SOV/NG languages are all NA. However this is the only area where there is a sharp difference: otherwise my figures are roughly half those of Hawkins' sample, bearing in mind the greater geographical bias of his and Greenberg's sample.

If we look at the languages in his sample which appear to show a significant pattern, we have

TABLE 3.8

SOV/Pr/NG/NA	SOV/Po/NG/NA
Persian	Sumerian
Iraqw (Cushitic)	Elamite
Khamti	Oromo ('Galla') (Cushitic)
Akkadian	Kanuri
	Teda
	Kamilaroi and other SE Australian languages.

Underlined are the languages also in my sample. Outside Hawkins' sample are related languages, such as Kurdish plus two Bantu languages, Tunen and Bandem, according to a footnote in Hawkins (1979), which are SOV/Pr/NG/NA. We thus have roughly eleven SOV/NG languages in the Hawkins sample, compared to my three, including Djingili (SOV/Po/NG/NA, a language from the Northern Territory of Australia).

To my knowledge however no languages have been found in the precluded groups, SOV/Pr/GN/NA and SOV/Po/GN/NA, except that perhaps Latin may display the first pattern, with dominant NA, and apparently GN for genitive of possession. If we disregard Latin, we may put forward the following points:

(i) The representation in my sample of one language for SOV/Pr/NG/NA and two for SOV/Po/NG/NA, and of course none for the unattested SOV groups is about right. A language in the sample should represent one in 4.75 sampling units and the above is close enough.

(ii) Statistics from three languages are likely to be somewhat less significant than those from eleven languages. This particularly applies when all the languages concerned all behave similarly, and for example place the adjective after the noun.

In this light we should pay special attention to the fact that the eleven SOV/NG languages in Hawkins' sample are all NA. However it emerges very strongly from the above tables that in Hawkins' sample as a whole, there are 63 NG languages, of which 50, or 79.4%, are NA. This provides us then with a very clear alternative to the Greenberg/Hawkins approach, precluding SOV/NG/AN languages, in that NG languages are overwhelmingly NA anyway, regardless of whether or not they are SOV. The question we should test then is whether eleven out of eleven is significantly different from 79.4 or roughly 80%. The answer is that we cannot really tell, since Hawkins's sample has not been systematically compiled, but the figure of around 80% is sufficiently high to give concern that we are really dealing with an exceptionless universal.

The above discussion has concentrated upon hypothesised tendencies among NG languages. I shall now consider GN languages, where the evidence of both samples is that an even distribution between NA and AN is appropriate. Hawkins' prediction is that Pr/GN languages will almost always be AN, upon the basis of the following data:

TABLE 3.9

	Pr/GN/NA	Pr/GN/AN
VSO	none	Milpa Alta Nahuatl
SVO	Arapesh	Swedish, Danish
SOV	none	Amharic

To these we may add Sulka (SVO/Pr/GN/NA) and Latvian (SVO/Pr/GN/AN). Taking the original data of Hawkins's sample, we have one NA language to four AN: is this significant? If we assumed a systematically selected sample there is approximately a one-in-six chance of this happening, and when the other two languages are added, it becomes approximately one-in-four (see binomial distribution, Hoel (1960)). Given that we are likely to be doing considerably more than six such tests in a work such as this, we can scarcely regard it as significant.

As a result of the above examination, it would seem that the exceptionless universals that Hawkins puts forward are rather less obviously motivated by the evidence, and that the results that he focuses upon should instead be taken as evidence of a correlation with genitival position, regardless of Pr/Po and verb position.

In addition to Hawkins I, II, III, and IV we have the following hypotheses:

V. $Pr \supset (NG \supset NR)$

IV. $Pr \supset (VO \supset NR)$

VII. $Pr \supset (NPoss \supset NA)$

and VIII. $Pr \supset (ND \supset NA)$

where NR stands for noun plus relative clause, NPoss for noun plus 'possessive adjective' and ND for noun plus demonstrative. Thus if a language is Pr and NG, it will also be NR, etc.

From these Hawkins formulates a Prepositional Noun Modifier Hierarchy. This we may see by taking the above universals in the following order:

VIII. Pr ⊃ (ND ⊃ NA)

II. Pr ⊃ (NA ⊃ NG)

V. Pr ⊃ (NG ⊃ NR)

Thus if we have a prepositional language that is ND, it will also by NA. If it is NA, then it will also be NG, and if it is NG we will also have NR. This gives us the following legitimate combinations for Pr Languages:

Pr	DN	AN	GN	RN	eg.	Amharic, Mandarin.
Pr	DN	AN	GN	NR		Scandinavian, Baltic.
Pr	DN	AN	NG	NR		German, Icelandic, Greek.
Pr	DN	NA	NG	NR		Italian, Mâsai.
Pr	ND	NA	NG	NR		Welsh.

But again we have exceptions:

Pr	ND	AN	NG	NR	eg.	Efik
Pr	DN	NA	GN	NR		Arapesh, Sulka

though I shall argue in my last chapter that there are very clear exceptional circumstances in the case of Efik and I shall thus sustain my hypothesis that AN and ND do not co-occur in natural language.

Greenberg himself in his 1963 paper observed that Efik showed this exceptional AN/ND pattern, exceptional indeed to his Universal 18:

> When the descriptive adjective precedes the noun, the demonstrative and the numeral, with overwhelmingly more than chance frequency, do likewise.

At this stage therefore we must question whether we are dealing with exceptionless universals, and in particular whether such multi-valued universals yield the advantages of exceptionlessness that he urges. For example can we not simply put forward three statistical patterns:

$$\text{Pr} \;\dot{\supset}\; \text{NG}$$
$$\text{NG} \;\dot{\supset}\; \text{ND}$$
$$\text{ND} \;\dot{\supset}\; \text{NA}$$

where the second two hold regardless of whether we are dealing with Pr languages or not?

We must therefore attempt to take a more general look at the pattern and from Hawkins' sample and my own, we see the following picture:

(i) (a) VSO/VOS languages are overwhelmingly Pr,
(b) SOV languages are overwhelmingly Po,
(c) SVO languages may be either Pr or Po, the former prevailing, with certain implications for voice, (and perhaps aspect, bearing in mind Ewe, Lugbara and Kashmiri).

(ii) (a) Pr languages are overwhelmingly NG,
(b) Po languages are overwhelmingly GN.

(iii) (a) NG languages are predominantly NA,
(b) GN languages are roughly equally NA and AN.

To this it may be possible to add on, from Hawkins' and Greenberg's work:

(iv) (a) NA languages may put demonstratives and numerals either before or after the noun - though the choice does not usually arise <u>within</u> the language - and may similarly place adverbs before or after the governing adjective,

(b) AN languages are overwhelmingly DN, NumN (numeral first), and Adv-Adj.

The first three sets of correlation have already been touched upon in earlier sections, when correlations were exemplified both between and within the major strata of my sample. We saw not only that Pr/Po showed a strong correlation with both VO/OV and NG/GN, but that such correlation could generally be detected within strata. This contrasted with NA/AN, which correlated with NG/GN only between strata, suggesting that such correlations were the net effect of long-term influences.

The question then arises of what mostly correlates with what: the conclusions listed above appear to take the form of a chain, and this is what I shall consider in detail in the next section: such a chain suggests that adjectival ordering may not only take the form of a statistical universal involving NG/GN to the exclusion of other variables, such as Pr/Po, or verb position. Thus my next section will concern itself with the formal properties of such a chain.

3.3 Arguments for the Prediction Chain

In this section I shall look further at the question of three-variable and multi-variable analysis, and examine the formal properties of what I shall term a prediction chain, reflecting patterns of conditional independence of two items, given the attributes of a third. I shall illustrate this here with respect to the orderings VSO/SVO/SOV, Pr/Po, NG/GN and NA/AN.

In section 3.2 the following three correlations were observed,

(i) between VSO/SVO/SOV and Pr/Po,

(ii) between Pr/Po and NG/GN,

(iii) between NG/GN and NA/AN

The question arises from such correlations, of whether there is any direct dependence between say NG/GN and VSO/SVO/SOV over and above that resulting from their respective correlations with Pr/Po. If not, these two items are conditionally independent, given Pr/Po. The question may be illustrated in tabular form: clearly if we can detect a correlation between VSO/SVO/SOV and Pr/Po

TABLE 3.10

	VSO	SVO	SOV
Pr	25	30	5
Po	1	15	60

and between Pr/Po and NG/GN

TABLE 3.11

	NG	GN
Pr	55	5
Po	7	69

- where the above data is taken from Hawkins' sample - we may well expect to find a correlation between VSO/SVO/SOV and NG/GN:

TABLE 3.12

	VSO	SVO	SOV
NG	24	27	11
GN	2	18	54

The correlation does indeed persist, although it is now weaker. The question then arises as to whether this latter correlation reflects some direct influence, or whether perhaps the influence is solely via its correlation with an intermediary, ie. Pr/Po.

We may seek an answer to this by setting up a multivariate table as follows, with the hypothesised intermediary always right at the top:

TABLE 3.13

	Pr			Po		
	NG	GN	Total	NG	GN	Total
VSO	24	1	25	0	1	1
SVO	27	3	30	0	15	15
SOV	4	1	5	7	53	60
Total	55	5	60	7	69	76

We can see that among Pr languages as a whole the ratio of NG to GN is roughly ten to one, and the reverse for Po languages. Now can we find any significant variation from this ten-to-one or one-to-ten pattern within VSO, SVO or SOV languages?

If we look at the above table we can see that these ten-to-one and one-to-ten patterns are not significantly deviated from. We may illustrate this by adding 'expected' frequencies:

TABLE 3.14

	Pr				Po			
	NG		GN		NG		GN	
VSO	24	(22.9)	1	(2.1)	0	(0.1)	1	(0.9)
SVO	27	(27.5)	3	(2.5)	0	(1.4)	15	(13.6)
SOV	4	(4.6)	1	(0.4)	7	(5.5)	53	(54.5)
Total	55	(55.0)	5	(5.0)	7	(7.0)	69	(69.0)

Here, for VSOPr languages, the expected NG:GN ratio is 23.0:2.0, almost identical to the overall Pr ratio of 56:5. In general we can see that the slight deviations of actual from expected cannot be attributed to the fact of a language being VSO, SVO, or SOV. In particular if we split the table into two, under Pr and Po, the resulting separate tables would not show anything that could be described as a statistical universal.

We see then that given Pr or given Po, there is no evidence from the data that verb position (VSO/SVO/SOV) and genitival position (NG/GN) are otherwise dependent. We may formally denote this by a chain of the form

$$\text{VSO/SVO/SOV} \longleftrightarrow \text{Pr/Po} \longleftrightarrow \text{NG/GN},$$

where the conditioning item Pr/Po, as intermediary, stands in the middle. It is important to note that as a result of this examination, we do not reject the hypothesis of conditional independence. We should note here that despite there being 137 languages or language groups in the above table (3.14), the result critically depends upon twelve languages - Pr/GN and Po/NG - in the middle two columns. To get a more certain result, it would be desirable to enlarge the sample - hopefully without any geographical or genetic bias - and employ a 'log-linear' statistical model (see Everitt (1979)). The reason why it is still worth-while pursuing the question despite the small number of 'inconsistent' languages is that the alternative orderings of the chain must be quite clearly rejected, and I shall show this now, by testing the alternative patterns, with VSO/SVO/SOV and NG/GN in the middle. The tests for these orderings are given in the following tables, where VSO/SVO/SOV and NG/GN are in turn treated as intermediaries:

TABLE 3.15

(A)	VSO		SVO		SOV	
	NG	GN	NG	GN	NG	GN
Pr	24	1	27	3	4	1
Po	0	1	0	15	7	53

and

TABLE 3.16

(B)	NG			GN		
	VSO	SVO	SOV	VSO	SVO	SOV
Pr	24	27	4	1	3	1
Po	0	0	7	1	15	53

Throughout this section I shall be considering a number of chains, which I shall term A, B, C, etc. We may note in Table 3.15, that there is a strong correlation between NG/GN and Pr/Po which cannot be explained by means of VSO/SVO/SOV as an intermediary. Under VSO, where 24 out of 26 languages are NG, it is highly unlikely that the only VSO language to be Po would be GN, unless there is some correlation between GN and Po. Throughout the tables for chains (A) and (B), one may find similar examples, where the peripheral items on the chain show a strong correlation with each other. This situation contrasts strongly with the situation where Pr/Po is the intermediary.

We thus accept the chain

$$\text{VSO/SVO} \longleftrightarrow \text{Pr/Po} \longleftrightarrow \text{NG/GN}$$

at least on the data available, and reject the two alternative orderings of the items on the chain. This represents the most straightforward

situation: that of accepting one ordering and rejecting the two alternatives. Throughout this work, however, other combinations - accepting two orderings, and rejecting all three - will also be encountered, and I shall defer any attempt to summarise their interpretation until some at least of them have been exemplified.

The next thing to consider is the addition of NA/AN to the chain. To this end we may first construct a master-table of the relevant data. The 'Greenberg numbers' are those assigned by Greenberg to his 24 combinations of word-order classifications (see Greenberg (1963)):

TABLE 3.17

	NG		GN		Greenberg numbers
	NA	AN	AN	NA	
VSO/Pr	19	5	1	0	1-4
VSO/Po	0	0	1	0	5-8
SVO/Pr	19	8	2	1	9-12
SVO/Po	0	0	6	9	13-16
SOV/Pr	4	0	1	0	17-20
SOV/Po	7	0	29	24	21-24

These figures may be compared with those listed in section 3.2: type 1 consists of 19 items, 15 languages and 4 language groups (VSO/Pr/NG/NA), type 2 consists of 4 languages and one language-group (VSO/Pr/NG/AN), etc.

Now the first thing we must test is whether NA/AN may play an intermediary role. To do this means looking at the following chains:

VSO/SVO/SOV ⟷ NA/AN ⟷ Pr/Po, (C)

Pr/Po ⟷ NA/AN ⟷ NG/GN (D)

Now the relevant tables, drawn from Table 3.17, are as follows:

TABLE 3.18

(C)	NA		AN	
	Pr	Po	Pr	Po
VSO	19	0	6	1
SVO	20	9	10	6
SOV	4	31	1	29

and

TABLE 3.19

(D)	NA		AN	
	NG	GN	NG	GN
Pr	42	1	13	4
Po	7	33	0	36

In neither of the above tables can the intervention of NA/AN account for the correlation between VSO/SVO/SOV and Pr/Po on the one hand, or between Pr/Po and NG/GN on the other.

We may now look at the possibilities of adding NA/AN at either end of the chain, as a peripheral, to give

NA/AN ⟷ VSO/SVO/SOV ⟷ Pr/Po ⟷ NG/GN,
or VSO/SVO/SOV ⟷ Pr/Po ⟷ NG/GN ⟷ NA/AN.

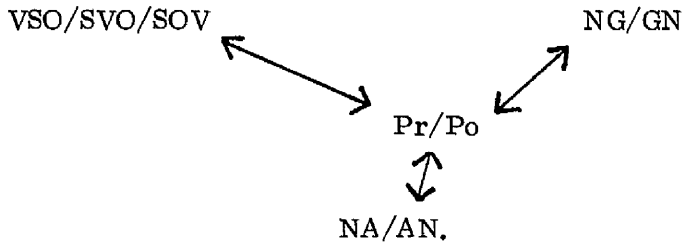
The first requires testing the sub-chain

NA/AN ⟷ VSO/SVO/SOV ⟷ Pr/Po, (E)

while the second requires

Pr/Po ⟷ NG/GN ⟷ NA/AN. (F)

However these are not the only possibilities: we also have the possibility that NA/AN is attached to the middle of the chain to give a 'three-legged chain' of the following form:



In this latter chain there are two sub-chains to test:

VSO/SVO/SOV ⟷ Pr/Po ⟷ NA/AN, (G)
and NG/GN ⟷ Pr/Po ⟷ NA/AN. (H)

The relevant tables are as follows:

TABLE 3.20

(E)	VSO				SVO				SOV			
	Pr		Po		Pr		Po		Pr		Po	
NA	19	(18.3)	0	(0.7)	20	(19.3)	9	(9.7)	4	(2.7)	31	(33.3)
AN	6	(6.7)	1	(0.3)	10	(10.7)	6	(5.3)	1	(2.3)	29	(27.7)

TABLE 3.21

(F)	NG				GN			
	NA		AN		NA		AN	
Pr	42	(43.5)	13	(11.5)	1	(2.3)	4	(2.7)
Po	7	(5.5)	0	(1.5)	33	(31.7)	36	(37.3)

TABLE 3.22

(G)	Pr				Po			
	NA		AN		NA		AN	
VSO	19	(17.9)	6	(7.1)	0	(0.5)	1	(0.5)
SVO	20	(21.5)	10	(8.5)	9	(7.9)	6	(7.1)
SOV	4	(3.6)	1	(1.4)	31	(31.6)	29	(28.4)

TABLE 3.23

(H)	Pr				Po			
	NG		GN		NG		GN	
NA	42	(39.4)	1	(3.6)	7	(3.7)	33	(36.3)
AN	13	(15.6)	4	(1.4)	0	(3.3)	36	(32.7)

It is noticeable that for tables (E), (F) and (G), the discrepancy between actual and expected varies from 0.3 to 1.5. However in table (H) the discrepancies are much greater, 2.6 under Pr, and 3.3 under Po. The sub-chain that this last table represents is part of the 'three-legged chain', and consequently this type of chain must be rejected for the above data.

As a result of these tests we may note that only the chain represented by (H) has been rejected: the three-legged chain. This conclusion is an extremely important one in that we have now rejected the idea that typological predictions are dependent upon one factor, such as VSO/SVO/SOV, or Pr/Po. Firstly we must reject VSO/SVO/SOV on account of the fact that any correlation between this factor and NG/GN may be explained by Pr/Po. This might suggest that Pr/Po is the one critical factor. Moreover even this is in doubt on account of the fact implicit in table 3.21 for chain (F), that any correlation between NA/AN and Pr/Po may be explained by NG/GN:

$$\text{Pr/Po} \longleftrightarrow \text{NG/GN} \longleftrightarrow \text{NA/AN.} \qquad \text{(F)}$$

Furthermore chain (H), of the form

$$\text{NA/AN} \longleftrightarrow \text{Pr/Po} \longleftrightarrow \text{NG/GN} \qquad \text{(H)}$$

was rejected on account of the above-mentioned correlation between NA/AN and NG/GN, as picked up in Table 3.23.

We now appear to have two conflicting alternatives open to us: firstly we may merge the chain that was originally accepted,

VSO/SVO/SOV ⟷ Pr/Po ⟷ NG/GN

with

Pr/Po ⟷ NG/GN ⟷ NA/AN (F)

to give

VSO/SVO/SOV ⟷ Pr/Po ⟷ NG/GN ⟷ NA/AN[5].

This is perfectly acceptable, since the alternative orderings of (F), with NA/AN as intermediary (D), and with Pr/Po as intermediary (H), have in fact been rejected.

However there is still a second alternative, obtainable by merging the first chain with

NA/AN ⟷ VSO/SVO/SOV ⟷ Pr/Po, (E)

in effect adding NA/AN at the other end, to give

NA/AN ⟷ VSO/SVO/SOV ⟷ Pr/Po ⟷ NG/GN.

Now we cannot reject this straight out, because chain (E) has not been rejected: Table 3.20 showed ~~that there was~~ no correlation between NA/AN and Pr/Po that could not be accounted for by their respective correlations with VSO/SVO/SOV.

The question therefore arises of how we deal with a situation where two alternative orderings of the four-element chain have not been rejected.

We must therefore ask whether the verbal position or the genitival position gives a better prediction for NA/AN. One possibility is to test the chains

$$\text{NA/AN} \longleftrightarrow \text{VSO/SVO/SOV} \longleftrightarrow \text{NG/GN} \qquad \text{(I)}$$

and

$$\text{VSO/SVO/SOV} \longleftrightarrow \text{NG/GN} \longleftrightarrow \text{NA/AN.} \qquad \text{(J)}$$

In effect we are examining whether NA/AN can be predicted by means of only one of the two other items. The relevant data is as follows:

TABLE 3.24

(I)

	VSO		SVO		SOV	
	NG	GN	NG	GN	NG	GN
NA	19 (17.5)	0 (1.5)	19 (17.4)	10 (11.6)	11 (5.9)	24 (29.1)
AN	5 (6.5)	2 (0.5)	8 (9.6)	8 (6.4)	0 (5.1)	30 (24.9)

TABLE 3.25

(J)

	NG		GN	
	NA	AN	NA	AN
VSO	19 (19.0)	5 (5.0)	0 (0.9)	2 (1.1)
SVO	19 (21.3)	8 (5.7)	10 (8.3)	8 (9.7)
SOV	11 (8.7)	0 (2.3)	24 (24.8)	30 (29.2)

Note that I have again shown in brackets the figures that would apply in each sub-table if no correlation occurs.

Table 3.24 shows very clearly that Chain I must be rejected: there is a very strong correlation between NG/GN and NA/AN among SOV languages. Thus there are no NG/AN languages whereas we would need something nearer five if the correlation between NG/GN and NA/AN could be accounted for by their respective correlation with verb position.

In Table 3.25 on the other hand the difference between the actual and bracketed figures is at most 2.3. Thus among the SOV/NG languages we find that all of these are NA, as Hawkins predicts, but this only differs by 2.3 from what we would expect if adjectival position could be predicted purely from genitival position.

Now while the difference of 2.3 is smaller than those encountered for the alternative orderings H and I it is nevertheless a larger figure than we have encountered elsewhere for chains that have either been accepted or not immediately rejected. (This example will be discussed further in a footnote to the next section.) It is important therefore in such analyses to ensure that the sample is truly representative. However we can see that if we are to predict NA/AN ordering in terms of a chain approach, we can only add it in accordance with chain J:

VSO/SVO/SOV $\longleftrightarrow$ Pr/Po $\longleftrightarrow$ NG/GN $\longleftrightarrow$ NA/AN

The above example serves to illustrate the concept of a prediction chain. It shows in effect that if we wish to predict the likelihood of a language being NA or AN, given a knowledge of its other characteristics VSO/SVO/SOV, Pr/Po and NG/GN, we need only make use of the last

one, NG/GN. This concept is particularly important in considering Hawkins' approach in which three items are joined in two implicational universals:

$$\text{Pr} \supset (\text{GN} \supset \text{AN}),$$
$$\text{Po} \supset (\text{NG} \supset \text{NA}).$$

In fact Pr/Po is not needed.

I have here deliberately chosen the Hawkins sample not merely to question his interpretation of his own data, but also because this particular sample, and its information on four major syntactic characteristics, is one of the best-known in the discussion of word-order universals over the last twenty years.

In taking this approach however it is dangerous policy to accept the sample while rejecting the conclusions. For all we know a more systematically compiled sample might sustain his approach. This is why it is important at each stage of the examination of statistical universals to make any improvements that appear necessary in the source data. My own choice of a sample therefore partly rejects the need to take steps to eliminate any genetic or geographic bias, and partly because a wider set of characteristics is necessary for the investigations in this work.

To illustrate the prediction chain, I have not only used Hawkins' sample, but also his choice of syntactic characteristics. For my own work, however, and the type of explanations I shall myself consider, 'trichotomies' such as VSO/SVO/SOV appear to blur the data not only on account of the under-estimation of VOS languages: in addition, if I am at any stage

to consider concepts such as operator-operand structures, it will be necessary to separate out VS/SV and VO/OV, and as the next section will show, the separate behaviour of these on the prediction chain is of particular interest.

3.4 Some difficulties with the prediction chain

In this section I shall look more closely at the prediction chain, and the problems associated with its use. To do so, however, I must separate out the influences of VS/SV and VO/OV, and introduce a concept of relative consistency, to relate VS with VO, Pr, etc. even though most VO languages appear to be SV. It is also time, having demonstrated the weaknesses of Hawkins's analysis of his own data, to examine results obtainable from my own data. In doing so I shall additionally consider problems of sampling which are illustrated by certain differences in the results obtainable from my sample and Hawkins's.

I shall proceed by first testing the chain

VO/OV ⟷ Pr/Po ⟷ NG/GN

and then

VS/SV ⟷ Pr/Po ⟷ NG/GN

TABLE 3.26a

Prepositional Languages

	NG	GN
VO	All Austronesian (10)	Mandarin
	Berber	Tiwi
	Cairene	SULKA
	Fulani	YUROK
	Greek	(4)
	Hausa	
	Kinyarwanda	
	Margi	
	Masai	
	Maung	
	Sango	
	Thai	
	Tiv	
	Vietnamese	
	Welsh	
	Yoruba	
	Zulu	
	CAYUVAVA	
	JACALTEC	
	SHUSWAP	
	ZAPOTEC	
	(30)	
OV	Iraqw	
	(1)	

Excluded: OLO (VO/Pr/NG~GN)

TABLE 3.26b

Postpositional Languages

	NG	GN
VO	-	Bardi (VSO)
		Ewe
		Finnish
		Lugbara
		GUARANI
		LUISENO
		ZOQUE
		(7)
OV	Djingili	Abkhaz
	Kanuri	Bawm
	(2)	Burmese
		Burushaski
		Hindi
		Ijo
		Japanese
		Kannada
		Mandingo
		Nama
		Newari
		Nubian
		Turkish
		Wati
		CARIB
		GUAYMI
		HIDATSA
		MAIDU
		NASIOI
		PIRO
		POMO (E)
		QUECHUA
		SIROI
Excluded: Ngandi (OV/Po/NG ⇝ GN)		All C. NEW GUINEAN (6)
		(29)

Individual languages are tabulated for the first chain in the two parts of Table 3.26. They may be summarised as follows, noting the exclusion of Ngandi and OLO:

TABLE 3.27

	Pr		Po	
	NG	GN	NG	GN
VO	30 (30.1)	4 (3.9)	- (0.4)	7 (6.6)
OV	1 (0.9)	- (0.1)	2 (1.6)	29 (29.4)
Total	31	4	2	36

Actual values and the bracketed values, calculated for no correlation within Pr and Po, are extremely close. We can see that all the correlation between VO/OV and NG/GN is explained by their respective correlations with Pr/Po. This bears out the results from Hawkins' sample.

Now the above table is certainly in accord with the results for the hypothesised chain from Hawkins' data: actual values and the bracketed values, calculated for no correlation within Pr and Po, are extremely close. But we are back to the situation in which very small numbers are involved: we have only one OV/Pr language in the present sample, compared with ten or so in Hawkins's, and from the present sample there is no way we can preclude the possibility, for example, that OV/Pr languages are evenly divided between NG and GN. There does indeed appear to be a preponderance of NG languages: Persian, Kurdish, Khamtî, Iraqw, Akkadian and the two Bantu languages Tunen and Bandem, against Amharic (GN) and Tigre and Tigrinya (both orders): seven to one with two varying. But these languages do not represent a systematically

sampled selection, and if we are to obtain such a selection, the likelihood is that we would have to make use of a much larger sample, perhaps of 300-400 languages.

Similar observations may be made about Po/NG languages: while it is certainly plausible that they split between OV and VO in a ratio of 4:1, as do Po/GN languages in the sample, we must bear in mind that with only two Po/NG languages in the sample, it is again possible that instead these divide evenly. It must be pointed out that all the Po/NG languages in the Hawkins sample are OV (except perhaps Moru-Madi, SVO/Po/NG ~ GN). But again this is not a systematically sampled selection.

The problem may be illustrated with respect to the three Semitic languages Amharic (OV/Pr/GN), Tigre and Tigrinya (likewise but with NG ordering also permitted). The question arises of whether the GN ordering is correlated with OV, or entirely independent: two chance inconsistencies with Pr. This is of particular interest since these are all three related to Ge'ez, which had VSO/Pr/NG ordering.

We see then the dangers of accepting a prediction chain ordering on the basis of a <u>lack</u> of correlation within each sub-table, where that lack of correlation is due to insufficient data. What we may observe however is that the sample does sustain the rejection of the alternative chains, partially at least:

$$\text{Pr/Po} \longleftrightarrow \text{VO/OV} \longleftrightarrow \text{NG/GN}$$

and

$$\text{VO/OV} \longleftrightarrow \text{NG/GN} \longleftrightarrow \text{Pr/Po.}$$

In the first case we may note that among the VO/Pr languages, 30 are NG languages to only 4 GN. On the other hand, the seven VO/Po languages are all GN. We may see this in Table 3.28:

TABLE 3.28

	VO		OV	
	NG	GN	NG	GN
Pr	30 (24.9)	4 (9.1)	1 (0.1)	0 (0.9)
Po	0 (5.1)	7 (1.9)	2 (2.9)	29 (28.1)
Fisher	0.00147%		9.38%	

Underneath each half of the table I have given the relevant probabilities that no correlation holds, employing Fisher's 'exact test'[6]. Thus there is approximately one chance in 68,000 that VO languages show no positive correlation between Pr/Po and NG/GN, a situation in which the bracketed figures would be more representative (at least if we rounded them).

We may draw up a similar table for NG/GN as intermediary:

TABLE 3.29

	NG		GN	
	Pr	Po	Pr	Po
VO	30 (27.2)	0 (1.8)	4 (1.1)	7 (9.9)
OV	1 (2.8)	2 (0.2)	0 (2.9)	29 (26.1)
Fisher	0.57%		0.36%	

Here the Fisher figures are both in the area of one in 150 to 300.

We must however turn back to Table 3.28, where the Fisher figure for OV languages was much higher, around one in 10. We may feel that this probability is too high, especially when we shall in this work examine rather more than ten such tables. In this case we cannot reject the possibility that among OV languages, NG/GN is directly influenced by Pr/Po.

It appears then that while the hypothesis of a Prediction Chain appeared to provide a very attractive alternative to Hawkins's search for exceptionless universals, and to yield a more satisfactory account than one which accepted exceptionless universals even though their statistical motivation was weak, a closer examination of the Prediction Chain has yielded difficulties which would suggest at first sight that a considerably larger sample is required. Furthermore we must note that though the Prediction Chain appears to provide an alternative to exceptionless universals, an exceptionless principle might in fact be inferred within the Prediction Chain concept itself: if for example we have the chain

$$\text{VO/OV} \longleftrightarrow \text{Pr/Po} \longleftrightarrow \text{NG/GN},$$

is the correlation between VO/OV and NG/GN <u>entirely</u> explained by Pr/Po, or are we simply saying in any particular test that a small correlation is possible, but that it doesn't show up from the sample? There is a difference between a small correlation and a non-significant correlation, and a larger sample might well reveal a small but significant direct correlation between VO/OV and NG/GN, perhaps exemplified by Amharic (VO/Pr/GN) and Tigre and Tigrinya (likewise but with NG $\sim$ GN). Any hypothesis which critically depends upon complete conditional independence between VO/OV and NG/GN given Pr/Po is highly vulnerable to refutation in the same way as Hawkins's approach. Therefore it is important, when examining Prediction Chain tests, to ensure that we are clear just what we are getting out of the results.

TABLE 3.30a

Prepositional Languages

	NG	GN
VS	Berber	-
	Cemuhi	
	Fijian	
	Maori	
	Masai	
	Niue	
	Tagalog	
	Tsou	
	Welsh	
	Yapese	
	CAYUVAVA	
	JACALTEC	
	SHUSWAP	
	ZAPOTEC	
	(14)	
SV	Cairene	Mandarin
	Chamorro	Tiwi
	Fulani	SULKA
	Greek	YUROK
	Hausa	(4)
	Iraqw (SOV)	
	Kinyarwanda	
	Malay	
	Margi	
	Maung	
	Palauan	
	Sango	
	Thai	
	Tiv	
	Vietnamese	
	Yoruba	Excluded: OLO (SV/Pr/NG ⇝ GN)
	Zulu	(Total Pr languages 36)
	(17)	

TABLE 3.30b

Postpositional Languages

	NG	GN
VS	-	Bardi (VSO) PIRO (OVS) (2)
SV	Djingili Kanuri (2)	Ewe Finnish Lugbara GUARANI LUISENO ZOQUE Plus all OV/Po/GN languages exc. PIRO (28) (see Table 3.26b) (34)

Excluded: Ngandi (SV/Po/NG ~ GN)

(Total Po languages 39)

Such problems may be further illustrated if we replace VO/OV by VS/SV, to complete our analysis of verb position. Relevant data is to be found in Table 3.30 and summarised in Table 3.31, where the ordering

$$\text{VS/SV} \longleftrightarrow \text{Pr/Po} \longleftrightarrow \text{NG/GN}$$

is tested:

TABLE 3.31

	Pr				Po			
	NG		GN		NG		GN	
VS	14	(12.4)	0	(1.6)	0	(0.1)	2	(1.9)
SV	17	(18.6)	4	(2.4)	2	(1.9)	34	(34.1)
Fisher	11.43%				-			

Under the Po column we have insufficient data to draw any conclusions, and under the Pr column we have an awkward figure of 11.43%, giving a one in ten chance approximately for independence between VS/SV and NG/GN among Pr languages: not a very large probability but large enough bearing in mind that rather more than ten tables are going to be tested. But the chief difficulty arises in that if the above chain is correct, with little or no conditional independence between VS/SV and NG/GN given Pr/Po, approximately four out of every 35 VS/Pr languages will be GN, just over 10%, ie the same proportion as for all Pr languages. Furthermore a good proportion of Pr/GN languages will be VS. But no such languages appear in the sample. This is not bad since there are only four Pr/GN languages in the sample. However there is only one language in the Hawkins sample with VS/Pr/GN ordering and that is Milpa Alta Nahuatl.

As a result there is some evidence for a direct influence, not only between VO/OV and NG/GN, as was conjectured for Amharic earlier in this section, but also between VS/SV and NG/GN. This leads us to try switching the ordering on the chain:

TABLE 3.32

	NG		GN	
	Pr	Po	Pr	Po
VS	14 (13.2)	0 (0.8)	0 (0.2)	2 (1.8)
SV	17 (17.8)	2 (1.2)	4 (3.8)	34 (34.2)
Fisher	32.39%		-	

Here NG/GN is the intermediary, and we see here, by contrast with the data in Table 3.29 where NG/GN is intermediary between VO/OV and Pr/Po, that the data for the chain

Pr/Po ⟷ NG/GN ⟷ VS/SV

is considerably less problematic. As a result the figures suggest a larger though problematic chain of the form:

VO/OV ⟷ Pr/Po ⟷ NG/GN ⟷ VS/SV.

Not only is there a problem here in that it would predict that Po/NG languages are roughly evenly divided between VS/SV, but also there is no sign of any representation of the fact that overwhelmingly we find that S precedes O. Now to deal with this problem I must turn to the concept of 'relative consistency': above we have a hypothesised chain, where each link between items has the form:

AB/BA ⟷ PQ/QP.

Each item represents a binary choice, unlike the one where VSO/SVO/SOV appeared when introducing the chain on the basis of Hawkins's data.

Now we could adopt the convention whereby all the items before the oblique are consistent with each other, and likewise after, as I have done informally already. But what does it mean to say that NG and VS are consistent, since in fact 19 of the NG languages are SV and only 14 are VS? We may therefore use the concept of 'relative consistency', whereby AB shows relative consistency with XY either if most XY languages are AB, or if AB forms a bigger proportion of XY than of YX (and hence BA forms a smaller proportion of XY than of YX). Thus VS languages, at 14 out of 33 NG languages, are a bigger proportion than one out of 40 is of GN languages. Thus we may safely term VS as relatively consistent with NG, and SV as relatively consistent with GN.

Now what does it mean if we have a chain of the form6a:

$$VO/OV \longleftrightarrow Pr/Po \longleftrightarrow NG/GN \longleftrightarrow VS/SV?$$

One result would appear to be that if we have a language such as Mandarin, with the order

VO/Pr*GN/SV

where the * marks a relative inconsistency, the language is unlikely to shift to VS unless it has first shifted to NG. Similarly it is unlikely to shift to OV unless we first have Po.

But why should such an ordering occur in the first place for the prediction chain? One possibility is that in SVO languages, (adverbial) prepositional phrases are more likely to follow the verb than precede it. The SV order

is more likely then to be incompatible with an NG order than with prepositional expressions, at least if genitive NPs are more common in subject NPs than PPs. It might also be thought that subjects and genitives are of a kind, whilst direct objects and pre-/postpositional objects are of another kind, but such an assertion would have to be substantiated by wider discussion.

It has been noted that if NG languages are split roughly 50-50 between VS and SV, then surely OV/NG languages would split 50-50 between VS and SV? We have three such languages in the sample: Iraqw (Pr), Djingili and Kanuri (Po), but all are SV. Furthermore examples outside the sample provide no strong evidence of any such VS languages and I shall return to this question - and exemplify it - in the next chapter when I consider incorporating case-marking into the statistics. We may however note that Iraqw permits VS ordering for intransitive verbs:

t huurín	ámêni	'The woman is cooking'
she is=cooking	woman	

In addition Djingili permits any ordering of subject, object and verb though SOV is clearly the dominant ordering. I have found no evidence for VS in Kanuri.

I do not wish therefore to rule out the possibility that Po/NG and OV/NG languages have a greater propensity to VS than is realised: alongside Akkadian (SOV/Pr/NG) we have Classical Arabic (VSO/Pr/NG). The NG construction of each is very similar - using a construct form for the head noun - and it is thus possible that Classical Arabic's VSO order was facilitated by a Pr/NG ordering. We also find that Kanuri (OV/Po/NG) and Moru-Madi (VO/Po/NG~GN) in the Nilo-Saharan family are related to the VSO language Masai.

Above is a survey of how VO/OV and VS/SV appear from the data to interact with Pr/Po and NG/GN. In view of the need to treat such chains with care when data is insufficient for part of the test tables, I shall now briefly summarise the areas in which the Prediction Chain is examined in this work for the purposes of theoretical evidence. There are two such areas.

The first area has already been exemplified in this section: of the four variables considered, VO/OV, Pr/Po and NG/GN showed very high correlations, with VS/SV showing a much weaker pattern. But since the actual ordering of items on the chain depends critically upon the rarer combinations such as OV/Pr, VO/Po, Pr/GN and Po/NG, results must be treated with considerable care. From a theoretical point of view, however, my main requirement is that these items correlate with an abstract concept such as operator-operand. One possibility is that non-subject NPs take part in such a structure, acting either as operator or operand to whatever governs them, the verb, pre-/postposition or possessed noun. But in this case there is no requirement that any one of these acts as intermediary on a Prediction Chain unless it most strongly adheres to this operator-operand structure: all we in fact require is a high level of inter-correlation. There is however limited evidence from this chapter that nothing other than Pr/Po, of the four constructions discussed, could be such a central item, and in my next chapter I shall further suggest that when inflectional case-marking occurs, this takes precedence over Pr/Po. None of this requires however that such a central item be the sole source of correlation, only a very strong source.

The second area in which Prediction Chains are relevant is that of adjectival modifiers. Here I attempt to assess the degree to which the semantic structure of these may be correlated with that of NPs and what governs them. This involves the attempt to add NA/AN to the Prediction Chain. Data is detailed in Tables 3.5 and 3.33. It turns out that any attempt to test the chain

$$\text{Pr/Po} \longleftrightarrow \text{NG/GN} \longleftrightarrow \text{NA/AN}$$

fails with my sample to determine the ordering of Pr/Po and NG/GN: all we can say is that NA/AN is not the intermediary. A larger sample may of course confirm the above ordering, as determined from the data of Hawkins's sample. There does however appear to be evidence for the chain

$$\text{NG/GN} \longleftrightarrow \text{ND/DN} \longleftrightarrow \text{NA/AN.}$$
$$(\underline{\text{or}}\ \text{Pr/Po})$$

This chain would appear to be a very convenient one, in that it conflates evidence that is already established, that NG languages are strongly NA, and ND languages almost exclusively so. As the evidence stands, neither of the other two items could act as intermediary, and the only evidence for a closed chain would come from the extremely rare AN/ND ordering. The possibility of such counter-evidence will be discussed in my last chapter, but apart from this possibility, it would appear that NA/AN has no direct correlation with Pr/Po other than via ND/DN.

3.5 The Trigger Chain

I now wish to look briefly at the consequences of combining stratal information with that for the prediction chains, and show that the resulting approach suggests that something like the so-called 'trigger chain' attributed to Lehmann and Vennemann is at work.

Essentially I shall replace Hawkins's 'Universal Consistency Hypothesis' with a 'High Odds' Principle, and a complementary 'Evens' Principle, each of which assume a prediction chain. Because I have identified the complementary 'Evens' Principle, I am able to replace the 'Universal Violation (or Trigger Chain) Hypothesis' with a more refined 'Diminishing Trigger Effect' which complements rather than conflicts with the two principles put forward.

Now if we take the chain

$$\text{VO/OV} \longleftrightarrow \text{Pr/Po} \longleftrightarrow \text{NG/GN},$$

we may add on data for the number of language families showing stratal variation (in brackets), and the number showing stratal correlation - above the arrows linking the correlated items:

$$\text{VO/OV} \overset{7}{\longleftrightarrow} \text{Pr/Po} \overset{5}{\longleftrightarrow} \text{NG/GN}$$
$$(9) \qquad\qquad (7) \qquad\qquad (6)$$

The hypothesis of the Trigger-Chain as formulated in Hawkins (1979), under the name of the 'Universal Violation Hypothesis', has already been introduced in §3.1, and is as follows:

> Languages violate synchronic implicational universals of the form 'If P, then Q' by evolving ~P and *Q co-occurences. Such co-occurrences trigger a chain of subsequent word-order changes which re-introduce consistency, as languages acquire whatever Q properties are observed to co-occur with P on syntactic evidence.

Hawkins argues against this on theoretical grounds: if change is towards restoring consistency, why should violations occur in the first place? There is in fact a fallacy in this argument, and we can see this in a non-linguistic trigger-chain, in the area of economics:

(i) If bad weather occurs, the availability of wheat drops.

(ii) If wheat availability falls, wheat prices go up.

(iii) If wheat prices go up, so does the price of bread.

(iv) If bread prices go up, so does the price of sandwiches.

This is undoubtedly a trigger chain, but there is no way we can argue against it by saying that as long as wheat prices are low, there is no reason why we shouldn't have good weather!

This is because the causation is one-way, since there is no trigger chain whereby we can get good weather simply by reducing the price of sandwiches, but as long as we have a situation in which good weather produces a drop in prices, we will have a symmetric correlation between weather and wheat prices. But the symmetry of the statistics does not mean that the causation goes both ways.

The distinction between symmetric statistical data and two-way causation is an important one. However we may find that even if the causation does not go two ways, we may still get a trigger chain, and here I come to the 'Diminishing Trigger Effect': we may see this by looking again at the prediction chain

VO/OV ⟷ Pr/Po ⟷ NG/GN

If we are predicting Pr/Po we may re-state Table 3.29 as follows:

TABLE 3.29 (re-stated)

	NG		GN	
	Pr	Po	Pr	Po
VO	30	0	4	7
OV	1	2	0	29

(Note that this is not the table for Pr/Po as intermediary, but the table in which Pr/Po and NG/GN are swopped around so that Pr/Po is on the far right-hand end.)

According to the above data we see that

(i) VO/NG languages are all Pr,
(ii) OV/GN languages are all Po,
(iii) VO/GN and OV/NG languages are more evenly distributed.

It would thus appear that if an item stands between two other items on the chain, and those two items are inconsistent with each other, eg VO/GN or OV/NG, there will be a comparatively even distribution between Pr and Po. We may put this forward as an 'Evens' Principle:

> When two syntactic structures are inconsistent, the status of any intermediary items on a prediction chain joining them will be subject to odds that are close to even.

It will appear that this principle holds as long as the correlations involved are fairly symmetric. The manifestation of such even tendencies, where roughly equal pulls prevail from opposite directions, will be discussed further in chapters 4 and 9. Indeed I shall give evidence that morphological fusion is the result of such conflicting pulls when one of the items in the ordering is a grammatical rather than lexical item.

I shall now turn to the situation where peripheral items on the chain are consistent. We may now obtain a 'High Odds' Principle:

> When two syntactic structures are consistent, there is a very high probability that any intermediary items on the prediction chain joining them will be consistent.

Not only are the odds extremely high, but the likelihood of getting no exceptions in a sample of say 100 languages is also strong. It is thus very easy to 'discover' Hawkins-style principles of consistency:

VO $\supset$ (Po $\supset$ GN),
OV $\supset$ (Pr $\supset$ NG).

Note that these principles are deduced solely from data in the sample. Outside the sample we have Amharic violating the second principle with OV/Pr/GN ordering.

But how, we may again ask, do such phenomena permit a trigger chain? To answer this, we may suppose another item on the chain, AB/BA, giving

AB/BA $\longleftrightarrow$ VO/OV $\longleftrightarrow$ Pr/Po $\longleftrightarrow$ NG/GN.

Now suppose that we have a language which is

AB * OV/Po/GN

where AB is inconsistent with OV and more relevantly Po. The 'Evens' Principle now applies. There is a roughly evens chance that the ordering will stay, and a roughly evens chance that it will change to

AB/VO * Po/GN.

If this happens the 'Evens' Principle may now apply again, but note that we are starting to creep rightwards along the chain, the probability approximately halving each time. There is now an evens chance of the structure changing to

AB/VO/Pr * GN.

Finally, if this change occurs, there will be an evens chance that the structure will change to

AB/VO/Pr/NG,

bringing about full consistency among those characteristics observed above.

We have seen in the above steps that the likelihood of a change from GN to NG is considerably less than say one from OV to VO, depending upon the cumulative effect of the 'Evens' Principle. Thus the impetus of the trigger dies out.

We may now look again at the chain

$$\underset{(9)}{\text{VO/OV}} \overset{7}{\longleftrightarrow} \underset{(7)}{\text{Pr/Po}} \overset{5}{\longleftrightarrow} \underset{(6)}{\text{NG/GN}}$$

and we may note that the data is entirely compatible with a situation in which VO/OV is most likely to change, with nine of the eleven language families in the sample with more than one language showing variation for VO/OV. However only six such families show variation for NG/GN, suggesting a diminishing impetus for the trigger chain.

We may represent this by a Diminishing Trigger Effect principle:

> If a language changes a syntactic construction represented in statistical terms on a prediction chain, such a change may or may not trigger changes along the chain. The further along the chain from the original change, the smaller will be the likelihood of such change, as a cumulative effect of the application of the 'Evens' Principle.

It must be pointed out though that the main evidence for the Diminishing Trigger Effect lies in the arguments concerning the application of the 'Evens' Principle. The diminishing figures 9-7-6 on the prediction chain above are by no means compelling on their own: they are merely not incompatible with the Diminishing Trigger Effect. More data would be desirable then to ensure that the figures 9-7-6 do in fact bear out a diminishing effect. This would require in particular getting data from the largest language groups represented in the present sample by only one language.

We can thus see firstly why illusory principles of consistency Hawkins-style are easily obtainable, and secondly that far from refuting the trigger-chain, once we have identified two compatible principles, the 'Evens' and 'High Odds' Principles, rather than two incompatible hypotheses, the Universals Consistency Hypothesis and the Universal Violation Hypothesis, the phenomena captured by the illusory principles of consistency actually support the concept of a trigger chain.

In this chapter I have attempted to demonstrate by example that there is no such thing as 'over-use' of statistical universals, where statistically significant patterns are to be found. Firstly I attempted to examine both exceptionless and statistical universals in terms of their use in identifying relevant definitional characteristics, in making predictions and in identifying phenomena requiring explanation, and here I showed that even though statistical universals were sufficient only in the third application, they formed a necessary element in the first two, in that any exceptionless universal that had a role to play in definition and prediction also needed to be statistically significant. Secondly I attempted to show that in a number of cases, the principles of consistency proposed by Hawkins, even though they represented valuable insights in themselves, were not statis tically significant, and that to be of further use they had to be re-formulated in a way that brought out significant patterns. It was here that the phenomenon of conditional independence became important, and for the purpose of the present study I have formulated the hypothesis of a network of prediction chains, in which correlations other than those represented by the chain were either non-existent or comparatively minor, not to be picked up without a considerably larger sample. The question now arises as to how such a chain is to be extended beyond what I have exemplified in this present chapter, and to what extent can the phenomena embodied in the chain be explained by semantic constraints, perhaps through the identification of operators and operands. To answer this I must look more closely at the ordering of verb, subject and object, and consider the role of case-marking, and the relevance of the 'Evens' and 'High Odds' Principles in accounting for morphological processes such as agglutination and morphological fusion. Such an investigation is useful not merely in filling out the prediction chain, but also in identifying the mechanisms at word level in the face of consistent and inconsistent word-order patterns.

TABLE 3.33

SOV Languages

	NA	AN
NG	Djingili	-
	Iraqw (Pr)	
	Kanuri	
	(3)	
GN	Abkhaz	Burushaski
	Bawm	Hindi
	Burmese	Ijo
	Mandingo	Japanese
	Nubian	Kannada
	Wati	Nama
	DAGA	Newari
	GUAYMI	Turkish
	HIDATSA	CARIB
	KATE	FORE
	NASIOI	MAIDU
	POMO (E)	MARIND
	SENTANI	QUECHUA
	WASKIA	SIROI
	(14)	(14)

Excluded: Ngandi (Po/NG ~GN/NA ~AN).

(Total SOV languages 32)

TABLE 3.34

VSO Languages

	NA	AN
NG	Berber	SHUSWAP
	Maori	(1)
	Masai	
	Niue	
	Welsh	
	Yapese	
	JACALTEC	
	ZAPOTEC	
	(8)	
GN	-	Bardi
		(1)

(Total VSO languages 10)

Note that the VOS languages Cemuhi, Fijian and CAYUVAVA are not included above. Nor are the verb-initial languages Tagalog and Tsou included, for which no dominant VSO or VOS ordering has been determined.

4. Agglutination, fusion and isolation

0.

In this chapter I wish to turn to the question of morpheme-order, where inflectional morphemes are identifiable and when the ordering of such morphemes is also identifiable. To this end I shall consider the interrelation between the tripartite typology of agglutination, fusion and isolation[1], and the more recent word order typologies, with their associated concepts of consistent and inconsistent ordering. In particular I shall consider isolation as the manifestation of consistency and fusion as a manifestation of inconsistency.

Traditionally the tripartite typology has been concerned with characterising languages, so that an isolating language will have one morpheme per word: this chapter however will be primarily concerned with the characterisation of constructions, and the idea that such a typology can characterise languages as a whole arises only when it appears from prediction chain analysis that one or more such constructions, in particular subject and object inflection, play a central role in such chains.

Evidence will be provided then that the traditional tripartite typology, initiated by the Schlegel brothers and further developed by Wilhelm von Humboldt, has a connection with the ordering typologies of Greenberg. However by asserting that adherence to ordering constraints is a necessary condition for isolation, and that violation of them is a sufficient condition for fusion, I am approaching the matter in a rather different way to that of Lehmann (1973), when he characterises SOV languages as agglutinating and VSO languages as fusional. What is more likely under my approach is that when suffixes are agglutinative in an SOV language, suffixes of a similar meaning will be fusional in VSO languages. But <u>prefixes</u> of similar meaning in VSO languages will be agglutinative. Such an approach initially requires

at least some concept of what the relevant ordering constraints are, but once the evidence is established for inconsistency as a motivation for fusion, the presence of fusional or isolating structures will provide further evidence in the continuing identification of such syntactic and morphological constraints.

An important task in identifying ordering constraints is the assessment of the degree to which agglutinative morphemes observe similar rules to analytic morphemes with similar meaning content. Thus in Turkish the word adamı is 'man' in the definite object case: -ı is the case suffix. In Hindi the corresponding expression is 'mard ko', where the definite object is indicated by the postposition ko, a separate word. Part of the statistics presented in this chapter will compare the behaviour of such synthetic and analytic elements. To do this, however, it will be necessary to distinguish between two types of non-inflectional constructions: those relying purely on word-order, if that, to identify functions such as case, and others which rely on analytic elements such as ko, and only differ from agglutinative constructions in the presence or absence of word-boundaries. It is important to note that in doing this we are going outside the bounds of morphosyntactic categories and their realisation: instead we must for each construction compared identify the additional syntactic and semantic criteria by which we determine what is to be compared.

In this and later chapters I shall show how the analysis of morphological patterns fits in with the prediction chain analysis.

4.1 Tripartite typology: a re-assessment

The tripartite typology as traditionally developed makes critical use of the word for the purpose of characterising languages. Thus a language may be treated as isolating if it employs no more than one morpheme per word. If however we treat a construction as isolating, the number of morphemes per word is no longer definitive: the internal form of the word is either constant or irrelevant, and the external form becomes critical, that is the place of the word within the containing phrase or clause. We can see this in a phrase such as 'John collects stamps'. Here we may consider 'stamps' as isolating for case, though not for number. For case the external form is critical: the ordering of the NP with respect to other items in the sentence. For a fusional or agglutinative structure on the other hand, the internal form of the word becomes critical, while the external form is either constant or irrelevant: thus English does not indicate plurality by means of word-order, though changes in word-order may occur for other reasons, such as topicalisation. We cannot however use case or number to characterise English as isolating or otherwise, since the two constructions make use of different devices.

To illustrate the distinctions further we may compare indirect object NPs in Turkish[2] (agglutinative), Latin (fusional) and English (isolating):

{ [adam- a] kitabı verdim } (Turkish)
man to book (Acc) gave (1sg)

{ [vir- ô] librum dedi } (Latin)
man Dat Sg book gave (1sg)

{ I gave [the man] the book. } (English)

Note here that I have used square brackets to enclose the internal structure and the curly brackets to enclose the relevant external structure: in the first two examples the relationship of the inner bracketed expression to the verb is indicated by the ending -a in Turkish and -ô in Latin, while in English the phrase 'the man' is Dative on account of its position in the clause. We may therefore consider agglutination and fusion to be morphological phenomena, and isolation to be governed by syntax. The contrast between Latin and Turkish may be elicited when we consider the plural counterparts:

{[adam-lar-a]$_{Pl}$ kitapları verdim.} (Turkish)

{[vir-îs] librôs dedî.} (Latin)

{I gave [the men] the books.}

Taking first the dative expressions, it will be seen that the suffix -a is unchanged: Turkish is thus agglutinative with respect to case and number, at least with respect to this particular case. The plural marker is in addition agglutinative with respect to the expression it pluralises, though this is not what we are comparing at the moment. Turning to Latin, we find that the -ô ending is no longer there: we have fusion of case and number, part at least of the expression put into the dative case. We may also argue that there is fusion with the stem, in that we cannot predict from either the nominative vir, or even the accusative virum that the dative forms will be virô and virîs. This situation requires more discussion, but it must be noted that in this instance, morphological rules must have access to the lexicon, rather than simply to a paradigm. This applies too to the English example, where number and stem are fused, at least in this example.

A similar analysis would apply to the accusative expression, where the Turkish example employs the definite object suffix -ı.

We have thus seen that Turkish is agglutinating for case and number, that Latin is likewise fusional, and that in the above example at least, English is isolating. But what happens when English makes use of the prepositional dative as is normal when the indirect object is focussed upon, and no special recourse is made to intonation (see Creider (1979)? To answer this we may look at the English sentence

$$\{\text{I gave the book } [_{a} \text{ to } [_{b} \text{ the man.}]]\}$$

and compare it with the Turkish

$$\{\text{kitabı}^{*} \; [_{a} \text{adam-} \; \text{a} \;] \; \text{verdim}\}$$

book (Acc) man to gave (1 sg)

I shall first consider brackets (a), whereby a comparison is being made between the two languages. We may firstly note that the order of indirect and direct object has now been reversed: according to Lewis (1967), any element which is to be emphasised may be placed immediately before the verb. I have therefore translated a focussed item in English by an emphasised item in Turkish, and it may well be that closer investigation would show that the focus and emphasis are one and the same thing in this context. The point is additionally worth note, in that Turkish does not alter the internal structure of an expression to give it emphasis in this way, so that we could say that Turkish is isolating for emphasis.

Turning now to the English example, we see that the indirect object expression, enclosed in brackets (a), is distinguished from the indirect object by means of an <u>analytic</u> construction: we add <u>to</u> to form a prepositional phrase. Note however that the NP, enclosed in brackets (b), is itself <u>isolating</u>, but with respect to its status as a prepositional object of <u>to</u>.

The real difference then between English and Turkish with respect to the above construction is that Turkish is synthetic (and indeed inflectional), whilst English is analytic. In each case the critical manifestation of the function of indirect object appears within square brackets, in contrast with an isolating structure, where the manifestation of the indirect object occurs in the structure outside the square brackets.

We thus see that if we are simply characterising a language as a whole, we only need to distinguish isolation as the non-inflectional characteristic: however if we are to look at individual constructions, it is necessary to sub-divide it further into isolating and analytic. In doing so, we may however ask why a type, here isolation, can be split into itself and something else. This is done to maintain the parallel between analysis at word-level and analysis at phrase level: a phrase or word consisting of a single morpheme can only be isolating with respect to its immediate environment. Here the concepts are identical, and in this way the link in terminology is maintained between the traditional analysis and an approach providing its motivation.

The outline I have given so far may be summarised as follows:

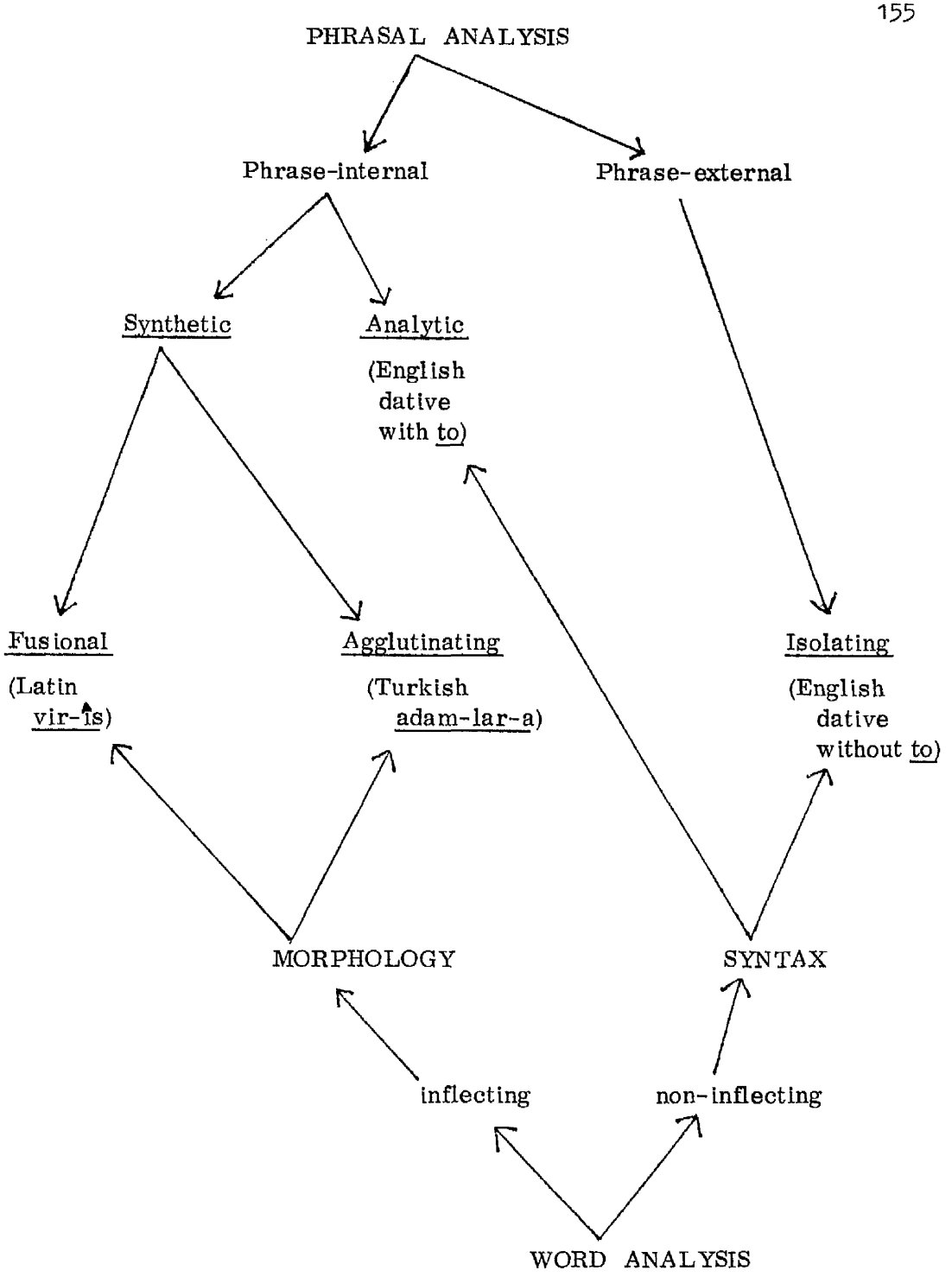

PHRASAL ANALYSIS
Phrase-internal
Phrase-external
Synthetic
Analytic
(English dative with to)
Fusional
(Latin vir-is)
Agglutinating
(Turkish adam-lar-a)
Isolating
(English dative without to)
MORPHOLOGY
SYNTAX
inflecting
non-inflecting
WORD ANALYSIS

Figure 4.1

It will be noted that my classification of 'analytic' differs from that of Sapir (1921). For him we may identify 'technique' (agglutinating, fusional/inflectional and isolating) as opposed to 'synthesis' (analytic, synthetic and polysynthetic). Thus for Sapir a construction may be both fusional and analytic at the same time. We may exemplify this in Latin:

vīsum	est,	it has been seen
seen	is	

This is both analytic, in the sense that we make use of the auxiliary esse, (to) be, and fusional, in the sense that vīsum fuses the elements of the lexical stem, perfect, participial, passive, neuter, singular and nominative, while est fuses the auxiliary stem, present, third person and singular, the latter showing agreement with the participle. But the point to be made is that the construction is only analytic plus fusional in the sense that the exponents are fused into separate bundles and the result joins analytically: if we only had two exponents, they could not join together analytically and fusionally at the same time. Consequently the terms 'fusional' and 'analytic' are not attributes of separate dimensions: they are of the same dimension, but each is used differently by the different dimensions inherent in the construction, ie the exponents.

4.1.1 Polysynthesis

We may make a similar statement in regard to polysynthesis if this involves the agglutination of lexical items: just as a single construction cannot be both analytic and fusional at the same time (see vīsum est in §4.1 above), there can be no construction which is both polysynthetic and fusional at the same time. Such a construction may be agglutinative for the synthesis of lexical items, but fusional for the synthesis of grammatical elements. The polysynthetism of Eskimo does not however involve agglutination of lexical items, but the agglutination of derivative morphemes (see Comrie (1981 b)), which correspond to separate lexical items in many other languages: such a concept of polysynthesis seems appropriate for the bilateral comparison of constructions.

4.2 The motivation for fusion

In this section I shall consider the motivation for morphological fusion. The hypothesis I shall put forward represents sufficient conditions for fusion, and is as follows:

(a) Agglutinative morphemes will follow the same ordering constraints as words (or 'analytic morphemes') in other languages with the same semantic content.

(b) When any ordering constraint is violated, a morpheme which is in the wrong place will fuse with all morphemes between this position and the position required by the ordering constraints.

(c) This fusion will be characterised by the fact that the meaning and form of the fused elements will be determined by lexicon and/or paradigm, rather than from the combination of the meaning of the separate elements.

It is important to note that the concept of fusion outlined in (c) is not necessarily confined to morphological structure: it may equally well apply to set expressions, consisting of two or more words. However such semantic fusion could only represent a single lexical item, since the combination of the meaning of the separate elements is precluded. Because the above hypothesis represents sufficient rather than necessary conditions for fusion however, there is nothing to stop set expressions arising out of a separate need for economy of expression.

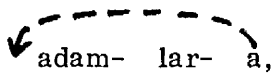
Fusional inflections for nouns have already been illustrated for Latin in the previous section, when

virīs (Latin)

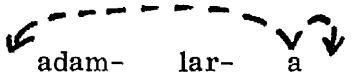
was contrasted with

adam+lar+a (Turkish),
to (the) men (English).

The hypothesis would mean that if for some reason the case-ending -a in Turkish became 'wrongly placed', in the sense that it would be expected to come initially:



then the -a would fuse with everything between itself and the empty slot pointed to by the arrow. It is important to note that the hypothesis does not say that something like this must have happened in some ancestor of Latin, since the hypothesis represents sufficient, rather than necessary conditions. If for example, VO/OV were the sole conditioning factor for the placement of case-marking (and later evidence will suggest that it is not), alternation of VO~OV might create a situation which may be represented by

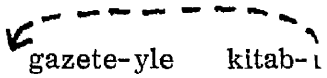


where marking is expected at both ends. In this case fusion might still be motivated, though this situation is not specifically considered in the hypothesis[3].

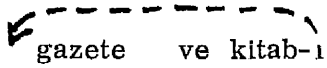
What happens when more than one lexical item stands in the way of the arrow? We may first look at co-ordinated expressions. There are three ways of saying in Turkish 'I read the newspaper and the book'[4]:

Gazete-	yle	kitab-ı	okudum	(Turkish)
paper	and	book Acc	read (1 sg)	
Gazete	ve	kitab-ı	okudum	
	and	Acc		
Gazete-yi	ve	kitab-ı	okudum	
Acc	and	Acc		

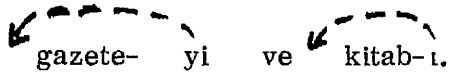
Thus we may co-ordinate either two NPs in the appropriate case, or alternatively co-ordinate the two NPs before adding case. Now if a situation arose in which case was required at the beginning rather than at the end, the first two examples



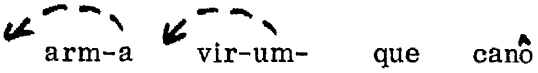
and



would require fusion of the whole expression, but not so when case is also co-ordinated:



A fusional construction resulting from the case-marking being in the wrong place would clearly require use of the last example's strategy to avoid fusion of the entire expression:

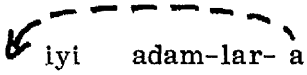


(Latin)

arm Acc	man Acc	and	sing
Pl	Sg		

'Of arms and (of) the man I sing'

A similar situation holds for adjective-plus-noun constructions. If again Turkish had an ordering violation



iyi adam-lar-a 'to (the) good men' (Turkish)
good man Pl to

the adjective would have to fuse with the noun to form a set expression.

In practice fusional languages repeat case-marking on noun and adjective. Not only do we see this among Indo-European languages such as Latin, Greek and the Slavonic languages:

tak- ûju bol'š- ûju rabót- u vy mne dáli (Russian)
such Acc big Acc work Acc you me gave
Fsg Fsg sg Nom Dat Pl

'You gave me such heavy work'

within the present sample all the languages displaying fusional case-marking also make use of case-agreement, with the exception of the prefixing language Berber.

If the position of these case forms is in fact in the wrong place, fusion of adjective and noun is not required: the case-marking on the first element is sufficient. However the question arises of how we treat the second and subsequent case/number inflections: either their correct position is straight before the item they inflect, or they are simply redundant, or as in the case of Turkish co-ordination, it is a matter of indifference.

We may exemplify this pattern with German, where we have examples such as

mein- em klein- en Freund
my DatSg little ? friend
SgM

'to my little friend'

Here it is only the first item that shows the distinctive case ending. The second item carries the weak declension inflection, which is -en for all except the nominative singular, and the non-masculine accusative singular. The first element shows the greatest variation for case in German, although the genitive and dative are neutralised for the feminine singular, and the nominative and accusative for all except the masculine singular.

In the examples above we see evidence that when we come across a stretch of utterance whose internal structure does not conform to ordering constraints, we cannot arrive at the meaning of the whole from the meaning of the parts. This is comparable in effect with Frege's Principle of Compositionality, in which the meaning of the whole is a function of the meaning of the parts and their mode of combination (see Dowty et al (1981)). Such ordering constraints then will apply to the mode of combination: fusion will occur and the Principle of Compositionality will cease to apply to a sequence of morphemes when such constraints are violated. By way of analogy we can say that if conditions for building a house are not suitable on the site itself, we must bring in pre-assembled material. The hypothesis then is that in the case of fusion the pre-assembly sites are the lexicon and the grammatical paradigms.

According to my hypothesis then, if case-marking is suffixal, yet the appropriate place is at the beginning of the noun phrase, it is necessary, to avoid fusion of noun and adjective, to make use either of case agreement, or of a system whereby the case suffix occurs only or with greatest contrast on the first element. The above examples have illustrated both alternatives - in Russian and German respectively. Data in the next section will show that when languages are both VO and Pr, the appropriate place for object marking, if it occurs, is at the front of the NP if fusion is not to occur, and that similar principles apply for subject marking.

4.3 Fusion and subject and object marking

In this section I shall consider the motivation for the fusion of subject and object marking on the Noun Phrase. In accordance with my hypothesis I shall try to show that when ordering patterns are adhered to, we may find agglutination and analyticity, where the meaning of the whole is determined from the combination of the parts, in this instance the subject/object marking and what is marked, but that when such patterns are not adhered to, fusion will occur.

Since I am including analytic as well as inflectional patterns in my study, it will be necessary to consider subject and object marking as more than just case inflections, which, as in Hindi, do not necessarily contrast subject and object. Instead I shall use case in a syntactic sense: of the 75 languages I have sampled, all but two, Tsou and Tagalog, appear to have either Accusative, Ergative or Tripartite systems (or a combination), and for the purpose of this section, case will be taken to mean the system which contrasts Nominative/Accusative, Ergative/Absolutive or for a Tripartite system 'Stative'/Ergative/Accusative, whether by means of word order, or segmental marking on the NP or verb.

Since object marking plays the same role as certain pre/postpositions in linking the NP to the verb, we may expect that non-zero object marking before the NP will correlate with prepositions, and like them correlate with NG and to some extent with VSO/VOS. Similarly we may expect such marking at the end of the NP will correlate with postpositions, and like them, with GN and to some extent with SOV. This is illustrated for the languages in the following table. Languages with inflectional or analytic marking for the accusative or absolutive case are shown, and the (a) denotes animate object, the (d) definite object and (opt) optional:

TABLE 4.1

ACCUSATIVE AND ABSOLUTE (Abs) CASE MARKING

	Pr		Po	
	prefix/Pr	suffix	suffix/Po	Form
VSO/NG	Maori			i/ki
	Niue (Abs)			e, a
	Welsh			soft mutation
	Berber			lexicon
SVO/NG		* Greek		paradigm
SVO/GN			* Finnish	lex/para
			GUARANI (a)	PE
			* LUISENO (a)	lexicon
SOV/NG		Iraqw		lexicon
			Kanuri (opt)	-gà
SOV/GN			Japanese (opt)	o
			MAIDU	-I
			Nama	-à
			Nubian	-Gâ
			QUECHUA	-ta
			Turkish (d)	-I
			(*) Hindi (d, a?)	para, ko
			HIDATSA (d)	-s
			Kannada (a)	-an (na)
			FORE (a)	nasal mutation
			POMO (E) (a)	-al
			Burmese (opt)	-Kou
			SENTANI (opt)	n[illegible]
			* Wati (Abs)	-PA, -NGA

The right-hand of Table 4.1 shows the form of the case-marking: only for agglutinative and analytic constructions are the actual forms given, with the agglutination shown by a hyphen. Items subject to morphophonemic processes are written with a capital. Case-agreement is indicated with an * , and the case-agreement patterns of Greek, Finnish, LUISENO, Hindi and Wati will be discussed in the text.

The most straightforward patterns to be observed are those of the VSO/Pr/NG and SOV/Po/GN languages. The first group uses initial marking, while the second much larger group has final marking, overwhelmingly agglutinative and suffixal. FORE is a rather perverse member of this suffixing group, in that it causes nasal mutation of the element following the object NP.

Fusion is indicated in the above table either by 'paradigm' or 'lexicon', the latter indicating that recourse has to be made to the lexical specifications for the item concerned. In effect we require principal parts.

It will be noted that of the 73 languages in the sample displaying case, only 24 appear in the above table. If however we tabulate the statistics of all the languages we get a very clear pattern:

TABLE 4.2

	VSO/VOS	SVOPr	SVOPo	SOV/OVS	Total
pref/Pr	4	0	0	0	4
isolating	9	20	3	17	49
suff/Po	-	1	3	16	20
Total	13	21	6	33	73

where the excluded two are Tagalog and Tsou.

In the above table we see that

(a) prefixal object marking goes exclusively with prepositional languages in this sample, while suffixing goes overwhelmingly with postpositionality,

(b) SVO prepositional languages are overwhelmingly isolating.

I shall look at the non-isolating data in detail, and we may note that among the 23 languages concerned, only a minority unconditionally mark the accusative or absolutive case. We may illustrate these examples from Maori and Nubian:

E	tiki	ana	ia	i	tetahi	kai	mana	(Maori)
Ipf	fetch	Ipf	he	Obj	some	food	for=him	

'He is fetching food for him (self)'

kitti	nassi	nulû	we-	kâ	udir-kâginî	(Nubian)
robe	long	white	a(n)	Obj	wearing	

'... wearing a long white robe'

Note that the -Gâ is realised as -kâ after a vowel.

By contrast Kanuri, SENTANI, Burmese and colloquial Japanese only optionally add a postposition or suffix. Thus alongside the Kanuri example

tɔ́ktɔ̀	kɔ́mu-	vè-	gɔ́	rúskɔnɔ̀
son	woman	Gen	Acc	I=have=seen

'I have seen the son of the woman'

we have

wú	kɔ́mbù	ɔ́tɔ̀	rúskɔnɔ́
I	blind=man	this	I=have=seen

'I have seen this blind man'

The Nubian and Kanuri examples also illustrate the tendency of agglutinative and analytic languages to add case-marking to the NP rather than to the noun: note in the Kanuri example how this results in the genitive and accusative suffixes clustering together.

We may additionally exemplify the optional use of case-marking from colloquial Japanese. According to Kuno (1972), the object marker o may be dropped:

boku	(wa)	kono	hon	(o)	katte
I	Top	this	book		bought

'I bought this book'

Note that the wa, concluding a topic expression, may also be dropped.

Wati and Niue each have Absolutive case-marking. Niue uses the prepositional particle e before common-noun NPs:

ne	kai	he	pusi	ia	e	moa
Pa	eat	Erg	cat	that	Abs	chicken

'That cat ate the chicken'

while the prepositional particle a is used before pronominal and proper-noun NPs.

Wati also varies the Absolutive case-marking, though this is placed on individual items in the NP, rather than at the periphery. Common nouns, adjectives and numerals only add an ending after a consonant, and this is -pa. Proper names, pronouns and demonstratives add -nga after a consonant and -nya after a vowel. In effect there is case-agreement for expanded Absolutive NPs:

wati-lu nganku-ku tjitji pika mankur-pa nga:-nya-ya katingu
man Erg me Gen child sick three Abs this Abs Pl brought

'The man brought these three sick children of mine'

As mentioned above, two languages, Tagalog and Tsou, have no specific case-marking as such, at least not to distinguish subject and object. This may be seen in the Tagalog examples:

bumbasa ng diyaryo ang titser
read newspaper Top teacher

'The teacher is reading a newspaper'

binabasa ng titser ang diyaryo
read teacher Top newspaper

'The teacher is reading the newspaper'

In the first example the infix -um- indicates that the verb is agent-focussed: the topic of the sentence is the agent. In the second sentence on the other hand, the verb is object-focussed and the topic of the sentence is the object. The ang precedes a topic, while ng precedes a non-topic subject, object or genitive, even in the same sentence. This concept of focus, as outlined in Schachter and Otanes (1972), is different from that discussed in chapter 1 of this work, but parallels a phenomenon outlined for Tsou by Tung (1964). The relevant morphology is partly effected by verbal inflection, and partly by a sentence-initial 'beginning part', which he outlines in his chapter on the topic[6].

The examples marked (a) in Table 4.1 essentially mark animate direct objects: these languages are GUARANI, LUISENO, Kannada, FORE and POMO (E). We may chart their practice as follows:

Pronouns	Propernames	Kinship	Human	Animate	Possessed
LUISENO	LUISENO	LUISENO	LUISENO	LUISENO	LUISENO
Kannada?	Kannada?	Kannada?	Kannada?	Kannada?	Kannada?
GUARANI	GUARANI	GUARANI	GUARANI	GUARANI	
FORE	FORE	(FORE)			
POMO(E)		POMO(E)			

Languages listed under each heading employ object-marking for the stated expressions. Hindi was marked (d, a?) and this will be discussed under (d).

In the case of Kannada the use of the accusative suffix -an(na) is optional, but it is most likely to be used with neuter object pronouns[7], then personal pronouns, then kinship nouns, then other nouns referring to people, and least likely with nouns referring to non-human objects. This hierarchy is stated in McCormack (1966), and the first item, neuter pronouns, would appear to violate the hierarchy of Silverstein (1976): however it should be noted that such pronouns may be omitted, so that the hierarchy is only partially violated.

A similar situation may be observed for GUARANI, where the postposition pe is used with reference to humans and sometimes animals and other entities[8]:

o-yuká	i-sý	pe	'he killed his mother'
he kill	his mother		
a-haɨhú	šẽ-retã́	mẽ	'I love my country'
I love	my country		

(note the change to mẽ after a nasal vowel)

but we have in contrast:

ho-'ú	so'ó	'he eats meat'
he eat	meat	

Object pronouns are normally omitted, being indicated by object inflection on the verb. When the pronoun is present, however, the postposition is obligatory, the form for the third person being šupé, giving (i)-šupé, him, her, it, them. We thus have further evidence for the tendency of inanimate pronouns to show a distinct object form, appearing to contradict the Silverstein hierarchy. It may be that this 'twist' to the hierarchy was overlooked in his analysis, due to the fact that Australian languages tend to be ergative.

As indicated above, LUISENO makes use of case agreement and the adjective is more likely than the noun to show object marking. I shall discuss LUISENO further at a later stage in this section, along with other languages showing fusional constructions.

The language FORE appears in the above chart, but kinship terms only take object marking if they are inalienably possessed. FORE also permits other nouns to take object marking, but this will be discussed below, along with the Tripartite EASTERN POMO, when the question of ergative marking will be examined. FORE may be further exemplified, on account of its unusual morphophonemic process, whereby each morpheme mutates the initial consonant of the next morpheme. The Accusative morpheme - when it occurs - has no phonemic realisation other than to trigger N-mutation in the following morpheme (or Q-mutation for pronominal objects):

Kabá:re	nk-ágaye	'He sees Kabare'
K.	N-sees	

Here the N-mutation adds nk- before a following vowel: in this example it is the first vowel of the verb, but in fact has equal effect on whatever follows.

Turkish uses the -I suffix for the definite object: the I is then subject to vowel harmony. Along with Hidatsa and Hindi it is thus marked (d). Hindi makes use of the postposition ko, the indirect object postposition, when the direct object is one which is 'individualised to some extent, and to which a degree of contextual importance is thus attached' according to McGregor (1977). He further observes that it will usually be used to refer to human beings and certain animals. But it is not used with pronouns which have a special accusative form, and like other postpositions it follows the Oblique case, which observes a fusional paradigm. For this reason we have the more complex entry 'para, ko' under 'Form'.

Examples have been given above in some detail in order to illustrate the problems we shall encounter when attempting to formulate ordering constraints upon case-marking. We may now turn to subject marking, and it will be noted that in my sample there are 23 languages with subject marking, a weighted percentage of 31%. The data is given in Table 4.3 where it will be seen that 13 of these languages are Ergative or Tripartite in some or all of their tenses.

TABLE 4.3

NOMINATIVE AND ERGATIVE CASE-MARKING

	Prepositional		Postpositional	
	prefix/Pr	other	suffix/Po	Form
VSO/GN			Bardi (Erg)	-nim
VSO/NG	Niue (Erg)			he/e
	Berber			lexicon
		*Mâsai (tone)		lexicon
SVO/NG		*Greek (suff)		paradigm
SVO/GN			*LUISENO	lexicon
			ZOQUE	-'Is
SOV/NG			Kanuri (opt)	-ye
			*Djingili (Erg)	-tu
SOV/NG ~ GN			*Ngandi (Erg)	gender
SOV/GN			Burmese (opt)	-Kà
			MAIDU	-Im
			Bawm (Erg)	nih
			NASIOI (Erg ?)	-(k)e
			Newari (Erg)	gender
			Wati (Erg)	-lu
			WASKIA (Erg i)	ke
			SIROI (Erg ? opt)	-nge
			POMO (E) (Tri)	gender
			Burushaski (Pf Erg)	-e
			GUAYMI (Pf Erg ?)	-we
			(*) Hindi (Pf Tri)	para, ne
			Japanese	ga

As with Table 4.1, the right-hand column shows the form of case-marking, with the actual forms given only for agglutinative and analytic forms, and the agglutinative forms distinguished by a hyphen. Items subject to morphophonemic processes are written with a capital and case-agreement is indicated by an *. The case-agreement patterns of Mâsai, Greek, LUISENO, Djingili, Ngandi and Hindi are discussed in the text. Note that Wati, which had case-agreement for Absolutive endings, has no agreement for the Ergative case-ending, which occurs at the end of its NP.

The data in Table 4.3 may be summarised as follows:

TABLE 4.4

	VSO/VOS/OVS	SVOPr	SVOPo	SOV	Total
pref/Pr	3	-	-	-	3
isolating	10	20	4	16	50
suff/Po	1	1	2	16	20
Total	14	21	6	32	73

where (i) PIRO is treated as OVS,
(ii) Tagalog and Tsou are excluded.

Of these 23 languages, twelve have not appeared in the table for object marking. Of these only Mâsai is clearly an accusative language. The remaining eleven are Bardi, ZOQUE, Djingili, Ngandi, Bawm, NASIOI, Newari, WASKIA, SIROI, Burushaski and GUAYMI.

The above analysis would suggest that Mâsai violates Greenberg's Universal 38:

> Where there is a case system, the only case which ever has only zero allomorphs is the one which includes among its meanings that of the subject of the intransitive verb.

The situation is in fact a little more complex, in that Mâsai uses tone to distinguish nominative and accusative:

édólít	olkítéŋ	'He sees the ox'
sees	ox (Acc)	
édólít	olkítèŋ	'The ox sees him'
	ox (Nom)	

To this extent it is difficult to determine which, if either, is basic. However the accusative is taken as the citation form, and in their Notes on Tonal Grammar, Tucker et al (1955) summarise tonal variation in the noun as if the nominative were derived from the accusative.

Accusative marking was observed above to occur more frequently with animate and definite objects. Ergative marking on the other hand was observed to occur more frequently with inanimate agents, and also with perfective verbs. We may illustrate the first with WASKIA and NASIOI. In these two languages the ergative marking corresponds to the instrumental marking and occurs with inanimate agents. We may note that each is classified as Indo-Pacific, though only WASKIA is Central New Guinean, and the suffix is ke in WASKIA and -(k)e in NASIOI[9].

buruk	ke	patete	naso	(WASKIA)
pig		sweet=potato	eats	

'The/a pig is eating sweet potatoes'

WASKIA also uses it optionally for human subjects, for both transitive and intransitive verbs, and examples in Rausch (1912) for NASIOI are consistent with a similar pattern for that language.

Table 4.3 above does not list FORE as having Ergative marking, although the suffix -wama is required for inanimate transitive subjects. Scott (1978), who does not analyse FORE as ergative, terms this suffix a 'delineator', and it may be added to any common noun in any case. It will be recalled that object marking is required for pronouns, propernames and certain kinship terms. When the delineator is added to common nouns they too require object marking, and Scott considers that the delineator serves to mark as a 'Potential Agent' something which would not otherwise be inherently so.

In EASTERN POMO the nominal subject of the transitive verb is marked by -he'- plus the third person pronoun, which varies for gender:

'íqanxa	da	he'-mit'	ba	šáhe'	ma·ráki·yàk^{h}·le
then	woman	+ she	that	fish	wanted

'Then the woman wanted that fish.'

This construction might be called the syntactic ergative: the suffixation of the subject pronoun is not required either for kinship terms or when an agentive suffix is already present:

ka·ló-	là·	wí	kó·qa	'A bee stung me'
bee	Ag	me	stung	

Agentive suffixes mark the semantic agent: they may thus be attached to something other than the syntactic subject of the verb. Additionally EASTERN POMO may mark the object, adding the -al to pronominal and kinship objects, and to objects marked by an agentive suffix as a semantic agent. Other nominal objects than kinship terms may optionally take the appropriate third person object pronoun, again linked by -he'-. Consequently we may regard EASTERN POMO as Tripartite, since we find both ergative and accusative case-marking but none for the intransitive subject[10].

Ngandi and Newari also have gender-sensitive ergative suffixes.

The use of the Ergative case with past and possibly future perfective verb-forms has already been noted. Hindi, Burushaski and apparently GUAYMI[11] exemplify this in the sample though in the case of GUAYMI this -we ending only applies to animate subjects. In the case of Burushaski and GUAYMI we have no special object marking for the non-ergative tenses. In Hindi on the other hand we do have such marking, the ko postposition described earlier this section, which may in fact be used in any tense. But we certainly cannot use ko for the subject of an intransitive verb, and therefore we may regard Hindi as Tripartite in the tenses that require ne for a transitive subject[12]:

aurat	ne	bacce	ko	bulayâ
woman	Erg	child	Obj	called

'The woman called the child'

baccâ	âyâ
child	came

'The child came'

In the above example we may note the paradigmatic variation in the noun baccâ, depending on whether or not it occurs before a postposition. Hindi is thus listed as 'para, ne', parallel to 'para, ko' for object marking. Above is a brief exemplification of ergativity as observed in my sample with respect to nominal inflections. Other languages, such as Chamorro and Abkhaz display ergative verb morphology. Others too, which have ergative or tripartite noun morphology, additionally make use of accusative systems for their pronouns, for their verb morphology or for both. In this section however I am primarily concerned with the behaviour of nominal case marking.

Having surveyed both subject and object marking I shall now turn and look more closely at the languages which determine case-marking by means of lexicon or paradigm. In this regard I shall first consider Iraqw, the one SOVPr example in the present sample. We may first note that certain nouns add the suffix -r to denote an object:

á	lákwânti-	r	leeleehíít
1sg	basket		looking

'I am looking for the basket'

though this suffix will also be used when followed by an adjective or a genitive NP. This class of nouns is in effect that of a syntactic feminine gender. The suffix is added to the head noun and not to the end of the NP. Case may also be indicated by a preverbal item, termed a 'Selector' in Whiteley (1958). This inflects for tense and subject agreement. It also inflects for object agreement if the object NP immediately precedes it. According to Whiteley, in his comments on word-order, this will occur if the object NP is to be given emphasis. Although we can state that object inflection on the noun is lexically determined in Iraqw, we can gauge little from the behaviour of only one SOVPr language in the sample.

For this reason we must examine such languages outside the sample. Such are Persian, Akkadian, Khamtî, Tigrinya and Amharic. Of these Persian and Khamtî place object marking, if any, at the end of the NP, Tigrinya places it at the beginning, and Akkadian and Amharic place it towards the beginning.

Looking first at Persian we find that the marker râ is used for definite objects:

hasan-é- pedar- é- ahmad- râ dîdand
Hassan é father é Ahmad Obj saw (3 pl)

'They saw Hassan, the father of Ahmad'.

man pûl gereftam
I money got (1 sg)
'I got (some) money'

There seems to be some indeterminacy in Persian orthography as to whether the râ should be treated as a suffix or a separate word, but Mace (1971) states that it is 'not a word, it is a particle, a suffix'.

Turning to Khamtî we find that according to Needham (1894) the item mai will follow the object:

man man mai au phai sật kâ
he village Instr fire destroy Past

'He burnt the village down' (lit. he destroyed the village with fire)

However examples may be found where the mai is omitted, in particular where the object NP ends with a numeral expression:

kau	nü	sâm	tô	yü	kâ	'I shot three deer'
I	deer	three	Cl	shoot	Pa	

or when mai occurs already in the clause to denote one of its other functions, those of marking dative, ablative or locative expressions:

phâ	nai	net	mai	wai	tâ	'Put the cloth in the sun'
cloth	the	sun		put	Imp	

(where nai, the, goes with what precedes).

In our other examples object marking is not NP-final: in Tigrinya the marking occurs initially[13]:

ə̀ti	xä̀lbi	nə-	naiti	sä̀b'ai	jubba	näxis – u
the	dog	Obj	of=the	man	coat	bit 3 sg

'The dog bit the man's coat'

In Akkadian and Amharic the situation is more complex, with suffixal object marking that occurs on the first rather than the last item of the NP. The Akkadian adjective shows agreement with the head noun in case, number and gender[14]:

maddatt-	a	ma'att-	a	Iû	amhur
tribute	Acc	much	Acc	indeed	received (1 sg)

'Much tribute I did indeed receive'

However the case ending may be lost before a genitive (and usually before a possessive suffix):

awat(-)	šarr-	im	ul	iṣṣuru
word	king	Gen	not	kept (3 pl)

'They did not keep the word of the king.'

However with the exception of an NP ending in a genitive noun, genitive pronominal suffix or relative clause, both the first and the last item in the NP will be marked for the case of the NP as a whole.

In Amharic too the earliest item will take the object marking, even if this is part of a preceding genitival NP:

yä-šum- u- n byêt ́ayyä
of chief the Obj house saw

'He saw the chief's house'

The evidence then from SOVPr languages is thus rather more mixed than with SOVPo languages: whereas in the latter languages, object marking was clearly suffixal when it occurred, here we have a variety of patterns, with the marking at the beginning, end or middle of the noun phrase. Sometimes case-agreement occurs, as in Akkadian, which also as it happens is fusional for case and number, while that of Amharic is agglutinative for case, but the marking goes back to the earliest element. The nə of Tigrinya is treated by Praetorius (1871) as a preposition, but in the orthography it is prefixed to the following word. It is thus unclear whether the construction should be treated as agglutinative or analytical.

Now we must look at VO languages. Here we shall also see that while VOPr languages tend to have initial object and/or subject marking, if any, the VOPo languages are rather more varied. In the present sample only GUARANI made use of object-marking that was clearly NP-final.

We may first look at Bardi, which is VSOPo, yielding an inconsistency between VS and VO, and Po. At first sight we might think that because ergativity is indicated by the suffix -nim, that case-marking, ergative

at least, opts here for consistency with Po. However we must note that the -nim, and apparently postpositions generally, have to be suffixed to the first item in the NP. We may see this in the following example:

agal	guluŋguluwara	iruluŋangal	nidi
and	Makabala=fruit	have=just=been=gathering	much

gayaṛa-	ɖini	-nim	bawa
'white'	group		children

'A group of 'white' children have been gathering large quantities of Makabala fruit' (Metcalfe (1975), part Ex 108)

Here the -nim is added to the first constituent of the NP concerned.

If we now look at Modern Greek, which has SVO and Pr, suffixing inflections for the direct object would appear to be inconsistent with both. However it is again a characteristic that the case-ending almost always occurs on the first element of the NP:

ékho	mián	adhelfí	ke	énan	adhelfó
I=have	a(n)	sister	and	a(n)	brother

'I have a sister and a brother'

It should be noted that the ending -n is only added to the article, adjectives and nouns if they are masculine or feminine singular, end in a vowel, and precede another word beginning with a vowel, k, p, t, ks or ps, at least in the same NP. For masculine plural nouns and adjectives ending in -i change this to -us. The article comes at the beginning of the NP and has its own paradigm, varying for case, number and gender. Now what is clear

here is that whatever the historical situation with the marking of the Greek accusative case, in modern Greek it is very much in evidence at the beginning of the NP, although in form it is essentially suffixing.

Thus we have in the Bardi ergative suffix and the Greek accusative endings a situation where in effect the suffixing has crept as far as possible to the beginning.

Turning to the Greek nominative case suffix, we have a situation where the suffixing is inconsistent with Pr, but is it relevant to regard it as inconsistent with VO, or consistent with SV? The only distinctive nominative suffix is the post-vocalic -s on masculine nouns, giving words such as anthropos, man, pateras, father, and erghatis, workman. It is never dropped. It also appears in the adjectival ending -os accompanying masculine singular nominative nouns, but not on the article o. Thus there is the appearance of a gender- (and number-) determined -s ending occurring towards the end of the NP, but certainly not creeping back in the way of the accusative marker.

It is probably unwise to give too much weight to the example of the Greek nominative -s, when the Greek declension is clearly paradigmatic. However, with ZOQUE, an SVOPo language, we have a clearly agglutinative -s ending for the Ergative case. It comes right at the end of the NP. As a suffixal form it is also consistent with Po and perhaps SV.

We may now look at three languages which are SVOPo and suffixing for the object marking. Here the suffixing is consistent with Po but not with VO. These are Finnish, LUISENO and GUARANI, which do however permit other orders for emphasis and other factors, including SOV.

GUARANI adds the postposition pe to animate NPs. There is no sign of fusion. With Finnish and LUISENO the object inflection takes the form of inflections on nouns and adjectives: provided that there is no preceding genitive noun phrase - both are GN - the case-marking when it occurs will occur on the first item in the noun phrase.

If we first take the case of Finnish inflections, we find a very irregular pattern for object marking. In general Finnish nouns require four principal parts:

(i) the nominative singular, eg kirkko, church,
(ii) the stem, eg kirkko-, from which all cases other than the nominative, accusative, genitive (sg) and partitive (sg) cases may be determined,
(iii) the genitive singular, eg kirkon, with the nom. pl. replacing -n by -t,
(iv) the partitive singular, eg kirkkoa.

Thus it appears that one cannot predict either the stem from the nominative singular, or vice versa. However Whitney (1956) provides some nineteen rules for forming the partitive from the stem - with exceptions - and a similar number for determining the genitive plural from the stem, even though this is not a principal part. The genitive singular (and hence the nominative plural) may also be obtained from the stem, by means of consonant-softening rules which are also complex.

Now the direct object is placed in the partitive if it is indefinite, or if the verb is either negative or imperfective. Otherwise the accusative case is formed by borrowing from the genitive in the singular or the nominative in the plural. Whether the borrowed genitive and partitive are obtained from principal parts, ie from the lexicon, or from the stem, ie

agglutination to it plus morphophonemic processes, we still require principal parts to distinguish it from the nominative singular. We thus have a situation predicted by the third part of my hypothesis.

I have indicated for LUISENO that both subject and object marking are determined by the lexicon: we find for example that some nouns add -a for the subject, others add -i for the object and still others do both or neither. Addition of -i is however the rule for adjectives and for plural and/or construct (ie possessed) nouns, as already discussed above.

It is characteristic of LUISENO that a noun or adjective in the nominative case non-construct form consists of a stem plus one of seven thematic elements or 'absolutive' suffixes (including zero), as in:

ya'a-	š,	man		ki:-	ča,	house
pa'qo-	t,	basket		to:-	ta,	stone
muvi-	l,	nose		hu:-	la,	arrow
	ano',	coyote				

where the 'coyote' example has a zero suffix, To form a plural an element involving -m- is added:

ki-	č-	um,	houses	to:- t- om,	stones

with some irregular patterns, including reduplication:

ya:yič-am, men

For the accusative, -a is dropped, when it appears on the absolutive suffix, and -i is added, though not necessarily with non-possessed substantives:

Nominative	Accusative	
ya'á- s	ya'á-č-i	man
ya:yi-č-am	yeyi-š-m-i	men
hu:-la	hu:-l	arrow
hu:-l-um	hu:-l-m-i	arrows
no-hu	no-hu- y	my arrow
no-hu:- m	no-hu:- m-i	my arrows

LUISENO resembles Kannada and GUARANI in that animates and pronouns tend to take the objective suffix. It differs, however, in that this appears to be a feature which has to be known for each lexical item, rather than a propensity within discourse for animates, etc. The use of the objective suffix on adjectives may be illustrated by the following object NPs (and by comparison Kannada makes use of no case-agreement):

AN:	muyuk-i	panál	'much yucca'
	much Obj	yucca	
NA:	kešmal	kehú:t-i	'small shell'
	shell	small Obj	

We may thus see that of the languages with an 'inconsistency' between VO and suffixation of object marking - Greek, Finnish, LUISENO and GUARANI - or an inconsistency between VS and subject marking - Bardi - we find that

(a) with the exception of GUARANI, the ending creeps back towards the beginning of the NP - indeed the one example of GUARANI parallels the exception of Tigrinya among SOVPr languages,

(b) the object endings used, if any, are conditioned by gender and/or number, where gender may include syntactic gender, as in Greek, or the semantic consideration of animacy, as in GUARANI,

(c) the endings of three of the languages, Greek, Finnish and LUISENO, are determined by lexicon or paradigm.

On the other hand, if we look at languages where prefixation is consistent with VO for object-marking and with VS for subject-marking, and languages where suffixation is consistent with SV for subject-marking and with OV for object-marking, we find that such marking will almost always be peripheral to the NP. There appear to be the following exceptions in the sample:

(a) Case agreement in three languages, Djingili, Ngandi and Wati, so that such case-marking is added not to the NP as such, but to nouns, adjectives and certain other items, as with Greek, Finnish and LUISENO above. We may note that each of these languages has alternate word-orders.

(b) Occasionally plural marking and numeral expressions are more peripheral to the NP than case and even pre- or postpositional marking. This has already been illustrated in the Wati sentence earlier in this section, exemplifying Absolutive case inflections. It is also to be found in ZOQUE:

pʌn-	hi'ŋ-	da'm	'with men'
man	with	Pl	

(see also Mandingo, chapter 6).

There are three languages, Berber, Mâsai (VSOPr) and Hindi (SOVPo) which show either case agreement or fusion, when this is not motivated by the hypothesis from the evidence of present-day languages. Mâsai indicates case by tone, and the case-number tonal paradigms are not at all simple. My hypothesis strictly does not preclude fusion occurring when ordering patterns are not violated - and indeed I have already discussed set expressions such as 'straight from the horse's mouth' where perhaps no violations appear - but unless such instances of unmotivated fusion are comparatively rare, my hypothesis would be difficult to substantiate. We must therefore say that since the other Nilo-Saharan languages in the sample are tonal and post-positional with no evidence of fusion, we have circumstantial evidence for irregular paradigms developing in Masai on becoming VO.

The example of Berber needs to be discussed carefully: each noun, unless it is invariant, has two forms to denote syntactic function. The object and pre-verbal NP make use of the citation form, whilst the other form, the forme d'annexion, is used when the noun is governed by an immediately preceding item, such as preposition, conjunction, numeral or preceding verb of which it is subject. The following examples show that knowledge is required from the lexicon:

Citation	Governed	(French)
argaz	urgaz	homme
aman	waman	eau
ifri	ifri	grotte
ils	yils	langue
tamazirt	tmazirt	pays
tagant	tagant	forêt
tigemmi	tgemmi	maison
tisent	tisent	sel

It is appropriate to list Berber in the above tables for subject and object marking as long as we may treat the pre-verbal NP as case-less, and say in effect that the pre-verbal NP is a pre-clausal item referred back to by an anaphoric pronoun, possibly implicit, within the clause. In this case the governed form includes subject among its functions, and the citation form object among its functions.

In Hindi fusional paradigms may be observed, not for ergative or direct object marking, but for two cases, direct and oblique, of which the latter occurs before postpositions and elsewhere as an instrumental or locative case. For many nouns the direct and oblique forms are identical, especially in the singular, and usually a postpositional phrase may be substituted for instrumental or locative uses. The inflectional system does in fact represent a heavily simplified set of paradigms since the days of Vedic and then Sanskrit. This erosion of the fusional pattern correlates with the overall drift remarked upon in Friedrich (1975), from a much weaker SOV pattern in Vedic, growing stronger in the prose of Classical Sanskrit, to a yet more regular albeit not completely rigid ordering in modern Indic languages, probably reflecting the influence of Dravidian.

We may also note the phenomenon of karmadhâraya, where adjective and noun are juxtaposed without case-agreement in Sanskrit, though this normally only applies to set expressions. Now this certainly fits into my hypothesis, where in the absence of case-agreement lexical items between the beginning and end of the noun-phrase will be fused. Since the word-order of Hindi is more strongly verb-final than Sanskrit, we have the example of a language becoming less fusional, with the resulting simplified paradigms (Direct vs Oblique) becoming candidates for a hypothesis concerning the transition from fusional to isolating languages.

The examples of case-agreement given above, where we see the case-agreement creep back in the NP, prompts to consider how common this is among languages which have VO⇝OV and VS⇝SV alternation. It is very difficult in fact to state categorically which languages do and which languages do not: thus Greenberg (1963) categorises Turkish as rigidly verb-final, but in the colloquial language post-verbal elements certainly occur, and such 'inversion' is known as devrik cumle (see Lewis (1967)). My understanding is that these post-verbal items cannot be extracted from a pre-verbal NP or subordinate clause, cannot bear emphasis and can only form an addendum to the intonation contour[15]. On the other hand, languages such as Nubian and QUECHUA, which certainly permit VO orderings, mark the accusative or ergative NPs peripherally. However, of the languages which show case-agreement for specifically accusative or ergative marking - Mâsai, Greek, Finnish, LUISENO, Djingili and Ngandi, all permit either VS⇝SV or VO⇝OV alternations or both. We may therefore express the above phenomenon in terms of the following implication:

Case-agreement (Erg/Acc) ⟹ Alternation (VS⇝SV, VO⇝OV)

In effect, since case-agreement does not occur without VO⇝OV and/or VS⇝SV alternation, the alternation must occur first, at least if the sample results are correct.

The data from the sample also provides strong evidence for another implication, although not an exceptionless implication. Of the languages with non-agglutinative subject/object inflection, Iraqw, Berber, Mâsai, Greek, Finnish and LUISENO, all but the first two make use of case-agreement[16]:

Fusion ⟹ Case-agreement

Above is a survey of the correlation of morphological fusion with word-order patterns, and it appears that we may look at the results from the point of view of two forms of inconsistency: the first is that between VO/OV and object marking - and of course between VS/SV and subject marking. Here the data suggests two stages:

(a) When VS⇝SV or VO⇝OV alternations occur, so that the position of subject and/or object marking is sometimes inconsistent, case-agreement may occur.

(b) When the predominant VS/SV or VO/OV ordering, with or without alternation, becomes inconsistent with the placing of case-marking, fusional processes become apparent.

We may note that (a) follows from the first of the two implications above, where case-agreement implies word-order alternation. It must be stressed however that data from prefixal constructions is insufficient: in addition it may be that fusion is motivated in stage (a), but it is only at stage (b), once the processes have taken effect, that an inconsistent VO/OV or VS/SV ordering becomes possible.

The second form of inconsistency that was noted was that of VO/Po and OV/Pr, where fusional object and subject marking were most noticeable. These strongly suggest that the 'Evens' Principle may be at work in some form, with case-marking standing between VO/OV and Pr/Po on the prediction chain, and perhaps something similar for VS/SV and either Pr/Po or NG/GN. To look at this possibility further, section 4.4 will consider the consequences of the above phenomena for the prediction chain.

4.4 The relevance of the prediction chain

So far I have considered the tripartite typology in terms of the motivation for fusion, with particular reference to subject and object marking. My basic hypothesis, stated at the start of §4.2, implies that word-order violations constitute sufficient conditions for fusion. In the last section however, we observed that though fusion was to be found far more with OV/Pr and VO/Po orderings than with other orderings, non-fusional constructions were certainly to be found as well: in particular the SV/Pr languages were overwhelmingly isolating. Consequently, if violations represent sufficient conditions for fusion, the precise form of the violations must still be found.

What we must now do is observe firstly that the 'sufficient condition' hypothesis does in fact represent the type of exceptionless universal discussed in the previous chapter (§3.1), where fusion represented the 'unless' situation. Secondly we may note that the presence of a wide distribution of inflectional patterns among VO/Po and OV/Pr languages suggests the application of the 'Evens' Hypothesis and hence the application of a prediction chain. I shall therefore devote this section to an examination of the data, with a view to establishing the relevance of such an analysis to the study of inflectional patterns. Having established the statistical patterns of fusion within a prediction chain account I shall then consider how such results can be made use of in an exceptionless hypothesis, since as I argued before, exceptionless universals should also be statistically significant.

I shall now turn to the first question, and as we shall see, the discussion of it will lead on to the second. I shall start with the testing of the prediction chain

VO/OV ⟷ Object marking ⟷ Pr/Po.

The consequences of this test I shall consider not only with respect to the chain

VS/SV ⟷ Subject marking ⟷ NG/GN

but also with respect to the chain

Pr/Po ⟷ NG/GN ⟷ VS/SV

for which there was some evidence in section 3.4 of the last chapter.

In the chain for object marking I shall distinguish four things:

(i) pre-marking (agglutinative or analytical),
(ii) isolating (or more strictly, uninflected),
(iii) fusional marking,
(iv) post-marking (agglutinative or analytic).

For ease of presentation I shall put the fusional figures in brackets after the isolating.

The table for the object-marking chain is as follows:

TABLE 4.5

	Object pre-marking		Isolating (and fusional)		Object post-marking	
	Pr	Po	Pr	Po	Pr	Po
VO	3	0	28 (2)	4 (2)	0	1
OV	0	0	0 (1)	17 (-)	0	15

(See Table 4.8 for full details: Tsou and Tagalog are excluded.)

Now the above pattern poses two problems: the first is the one I stressed in chapter 3 (§3.4), that the closer the correlation between three items, the fewer the exceptions and the greater the sensitivity of the prediction chain to these few languages. Indeed all agglutinative pre-marking languages in the sample are Pr, and all agglutinative post-marking languages in the sample are Po (though we have Persian and possibly Khamti as Pr languages outside the sample). We see this if we look again at the side columns in the above table:

TABLE 4.5 (extract)

	Object pre-marking		Object post-marking	
	Pr	Po	Pr	Po
VO	2	0	0	1
OV	0	0	0	15

But this brings us to our second problem, namely that the above pattern is of a type we have not encountered before: two-thirds of the languages are isolating with respect to the verb and its object. For these the extra link in the chain is not relevant, and for these, summarised in the middle columns of Table 4.5, a strong correlation between VO/OV and Pr/Po is observed. This would not normally be acceptable for a prediction chain: the chain would then be refuted and the chain possibly re-ordered.

Such an alternative ordering would be as follows:

VO/OV ⟷ Pr/Po ⟷ Object marking

with Pr/Po as intermediary. The consequences of such a hypothesis

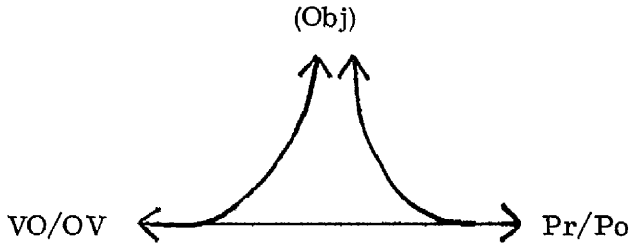
is that VO/OV and the placing of object marking will have no correlation save for their respective correlations with Pr/Po. This cannot however be the case, since as we have seen, VO/Pr languages are predominantly isolating, while OV/Pr languages have a variety of patterns.

We may alternatively look at the ordering of

Pr/Po ⟷ VO/OV ⟷ Object marking

But similar objections apply, since VO/Pr languages tend to be isolating or prefixal, while VO/Po languages tend to be isolating, fusional (Finnish and LUISENO) or suffixal (GUARANI). We thus have strong evidence from OV/Pr and VO/Po languages that the 'Evens' Principle (§3.5) is applying, with object marking in the middle of the chain. The VO/Pr and OV/Po languages provide similar evidence for the 'High Odds' Principle, (also §3.5), at least for those languages showing object inflection.

To capture this type of situation, we must depict it on the chain in a different fashion from what we have so far encountered:



With so many languages having an isolating pattern for the object NP, we clearly cannot say that all correlation between VO/OV and Pr/Po is via object marking. Therefore arrows link them directly. On the other hand, when object marking occurs, it is clearly affected by both

VO/OV and by Pr/Po. Therefore additional arrows are added, merging in with the others from the bracketed Obj. The merging of arrows is necessary to indicate that we do not have a closed chain when object-marking is present.

The above discussion refines the evidence on fusion from case-marking although subject-marking has still to be considered, in terms of the prediction chain. What we may note however is that the prediction chain evidence does not strictly back the idea that inconsistency is a sufficient condition for fusional processes - after all, GUARANI (VO/Po) and Tigrinya (OV/Pr) display analytic or agglutinative marking. Rather, it provides evidence that fusion is likely to occur as a result of the 'Evens' Principle applying when one or more of the items ordered is a grammatical item rather than a lexical item. Because of languages like GUARANI and Tigrinya, we can only show that fusion is a sufficient condition if we can show that for some reason, perhaps lacking a passive for GUARANI, that ordering constraints are not violated. This is of course something that requires further investigation, and in particular, identifying fully the semantic constraints.

What we can do at this stage with the evidence is exemplify the consequences of the prediction chain analysis for hypotheses such as Vennemann's of a shift from OV to VO via TVX as a result of phonological erosion of case-marking (see Vennemann (1974)). There are two problems about this hypothesis. One is that if VO comes about to avoid ambiguity through loss of case-marking, TVX will not help. The other is that in the case of the development of SVO among the modern Germanic and Romance languages, the original SOV languages were prepositional, so that the inconsistent ordering provided a motivation anyway, perhaps through the Trigger-Chain mechanism that Vennemann himself provides.

Now what we can see from the prediction chain analysis, given the chain

$$VO/OV \longleftrightarrow (Obj) \longleftrightarrow Pr/Po$$

for case languages, is that as long as object-marking is fusional, there need be no inconsistency with either Pr/Po or VO/OV. Nor need there be any direct inconsistency between Pr/Po and VO/OV since object marking intervenes on the prediction chain. Consequently as long as there is no erosion of case-marking an OV/Pr ordering will be stable. However should such erosion occur, we have no need to have recourse to the argument that SVO is required to avoid ambiguity, since we now have a direct inconsistency between Pr and OV. In this situation, depending on how the prediction chain extends beyond VO/OV, we will see either the application of the 'Evens' Principle, and there will be a 50-50 chance or thereabouts that OV will change to VO, or the application of the 'High Odds' Principle, in which a shift to VO will occur with a very high probability.

In the above discussion I have presented the evidence for treating object marking as an intermediary on the prediction chain between VO/OV and Pr/Po, discussed the additional problems resulting from the many languages which have no object marking, where direct dependence appears to occur between VO/OV and Pr/Po, and finally illustrated this somewhat complex phenomenon with respect to case erosion and word-order change. It is imperative however to bear in mind the problems of trying to set up a chain where there are only a few if any languages for certain of the inconsistent patterns. One must therefore emphasise that such a pattern is tentative: until we get more data on prepositional object-suffixing languages such as Persian and postpositional object-prefixing languages (no examples), there is always the possibility of three-way correlations and a closed chain. We can

see this if we look at possible postpositional object-prefixing languages: according to the prediction chain prediction these should be VO, with object-prefixation a random manifestation of the 'Evens' Principle. But if they were found to be overwhelmingly OV, we should have evidence for a three-way pattern of correlation and a closed chain. Another possibility is that postpositional languages never make use of object-prefixing: in this case our evidence for chain ordering would depend critically upon languages such as Persian.

I shall now turn to the question of subject-marking. For this purpose we may set up a table very similar to Table 4.5:

TABLE 4.6

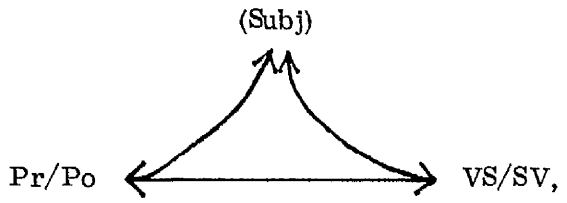
	Subject pre-marking		Isolating (and fusional)		Subject post-marking	
	Pr	Po	Pr	Po	Pr	Po
VS	1	0	9 (2)	1 (-)	0	1
SV	0	0	21 (1)	19 (1)	0	17

(See Table 4.9 for full details: Tsou and Tagalog are excluded.)

The pattern of this table is extremely close to that of Table 4.5 except that there are many more SV/Pr languages than OV/Pr. Can we therefore treat subject-marking as an intermediary between VS/SV and Pr/Po in the same way as object-marking appeared to act as intermediary between Pr/Po and VO/OV? The argument requires that VS/Po languages and SV/Pr languages will display a variety of patterns for subject-marking, in accordance with the 'Evens' Principle. In fact the data is extremely inconclusive. Our only VS/Po languages are

Bardi and PIRO, scarcely enough to detect a variety, while the 22 SV/Pr languages, (all SVO except Iraqw) are isolating with exception of the fusional suffixing of Greek. On the other hand the evidence from VS/Pr and SV/Po languages is clearer, expecially for the latter, where there are a large number of suffixing languages.

We thus have limited evidence for a chain of the form



and it is quite clear that we would require a rather larger sample to confirm such a pattern. Apart from anything else, very similar results would arise if we tested the chains

VO/OV ⟷ (Subj) ⟷ Pr/Po,

and VS/SV ⟷ (Obj) ⟷ Pr/Po,

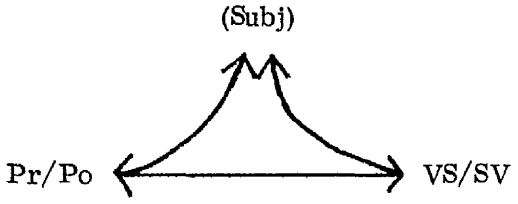
in which we compare subject-marking with object placing and vice-versa. This is because such subject and object marking occurs primarily in VSO, VOS and SOV languages where VO/OV and VS/SV are going to be consistent with each other in contrast with SVO (and OVS) languages.

However this is not the end of our examination of subject-marking, since we must now recall that VS/SV appeared to have a closer correlation with NG/GN than with Pr/Po. It is more revealing then at this stage to take into account the chain:

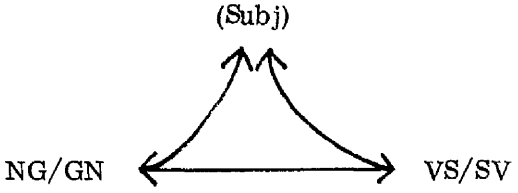
Pr/Po ⟷ NG/GN ⟷ VS/SV

which has so far been only tentatively established.

The problem is one of how we can reconcile the above chain with the pattern



The answer would appear to be that we have something of the following form:



where subject marking acts as an intermediary between VS/SV and NG/GN. The link with Pr/Po may then be effected either via NG/GN or via the close correlation between subject and object marking. Data for this link is summarised in Table 4.10, and may be summarised as follows:

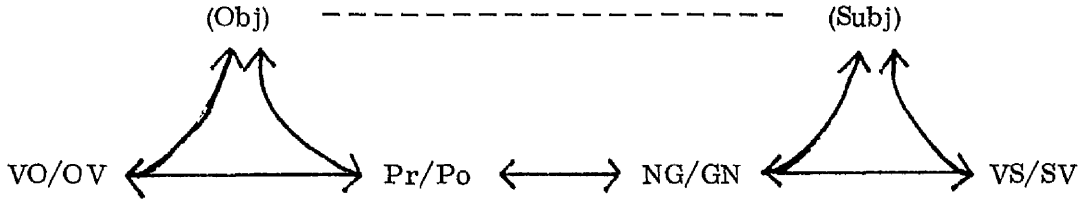
TABLE 4.7

	Subject pre-marking		Isolation (and fusion)		Subject post-marking	
	NG	GN	NG	GN	NG	GN
VS	1	0	9 (2)	1 (-)	0	1
SV	0	0	16 (1)	23 (1)	2	14

Note that not only Tsou and Tagalog are excluded from this Table, but also Ngandi and OLO, for which no dominant NG/GN ordering has been established.

In the above table we may first of all note the direct correlation between VS/SV and NG/GN: in particular there is only one VS/GN language in the sample, that is isolating for subject marking. Under the peripheral columns however, such a correlation cannot be discerned, partly of course because of the limited amount of data. But can we preclude alternative orderings of the chain? The limited evidence suggests that we must preclude NG/GN from being the intermediary: among NG languages the only one to use non-fusional pre-marking for subject is Niue, a VS language, while the only ones to use non-fusional post-marking are Djingili and Ngandi, SV languages. Similarly we have limited evidence that we must preclude VS/SV from being intermediary: among VS languages Niue is subject pre-marking and NG, while Bardi is subject post-marking and GN.

Now what this gives us is a structure of the form



where the respective influences of the Obj and the Subj are in fact difficult to separate. It is important to note however that the data for the right-hand end is extremely tenuous, particularly when the evidence for Bardi being GN is restricted, and more data is definitely necessary to be certain of the above conclusions.

The above structure gives us in a diagrammatic form the most likely relationship between word-order and morpheme-order patterns in terms of the data available. In effect they provide a minimum set of direct correlations, which may nevertheless have to be added to from a larger

sample, with more of the rarer inconsistent patterns, such as VS/GN and OV/Pr.

The structure suggests that when a language is isolating in terms of the subject and object constructions, the prediction chain represented by the word-order correlations is appropriate. When inflection does occur however, it is clear that subject and object marking plays a much more central role, even though their separate influences are difficult to discern.

What does the above analysis tell us about our hypothesis that inconsistent ordering is a sufficient condition for fusion? The point to be made is that because of languages such as GUARANI (VO/Po) and Tigrinya (OV/Pr) with analytic patterns, the nature of the relevant constraints has yet to be identified. However we may again observe that if fusion represents the 'unless' situation to an otherwise exceptionless universal, it must be statistically significant if we are to avoid the phenomenon of an exceptionless but statistically non-significant universal.

The fact that fusion occurs considerably more with VO/Po and OV/Pr languages than with other patterns suggests that we have found an 'unless' situation, without yet being able to establish precisely what it is an exception to. The time has come therefore to look more closely at the nature of the semantic constraints, and this I shall now do, firstly by examining the evidence available from the study of adjectives.

Table 4.8a

Languages with object pre-marking

(non-fusional)

	Pr	Po
VO	Maori Niue Welsh (3)	-
OV	-	-

(Total languages 3)

Table 4.8b

Languages with fusional object-marking

	Pr	Po
VO	Berber Greek (2)	Finnish LUISENO (2)
OV	Iraqw (1)	-

(Total languages 5)

Table 4.8c

Languages with isolation for object

	Pr	Po
VO	Cairene	Bardi (VSO)
	Cemuhi	Ewe
	Chamorro	Lugbara
	Fijian	ZOQUE
	Fulani	(4)
	Hausa	
	Kinyarwanda	
	Malay	
	Mandarin	
	Margi	
	Mâsai	
	Maung	
	Palauan	
	Sango	
	Thai	
	Tiv	Abkhaz
	Tiwi	Bawm
	Vietnamese	Burushaski
	Yapese	Djingili
	Yoruba	Ijo
	Zulu	Mandingo
	CAYUVAVA	Newari
	JACALTEC	Ngandi
	OLO	CARIB
	SHUSWAP	DAGA
	SULKA	GUAYMI
	YUROK	KATE
	ZAPOTEC	MARIND
	(28)	NASIOI
		PIRO (OVS?)
OV	-	SIROI
		WASKIA
		(17)

(Total languages 49)

Table 4.8d

Languages with post-marking for object
(non-fusional)

	Pr	Po
VO	-	GUARANI (1)
OV	-	Burmese Hindi Japanese Kannada Kanuri Nama Nubian Turkish Wati FORE HIDATSA MAIDU POMO(E) QUECHUA SENTANI (15)

(Total languages 16)

Note: data in this table may be cross-checked against data in Tables 3.4 and 3.6, bearing in mind that Tagalog and Tsou (each VO/Pr) do not appear in Table 4.8, for reasons discussed in the text.

Table 4.9a

Languages with pre-marking of subject
(non-fusional)

	Pr	Po
VS	Niue (1)	-
SV	-	-

(Total languages: 1)

Table 4.9b

Languages with fusional subject-marking

	Pr	Po
VS	Berber Mâsai (2)	-
SV	Greek (1)	LUISENO (1)

(Total languages 4)

Table 4.9c

Languages with isolation for subject

	Pr	Po
VS	Cemuhi	PIRO
	Fijian	(1)
	Maori	
	Welsh	
	Yapese	
	CAYUVAVA	
	JACALTEC	
	SHUSWAP	
	ZAPOTEC	
	(9)	
SV	Cairene	Abkhaz
	Chamorro	Ewe
	Fulani	Finnish
	Hausa	Ijo
	Iraqw(SOV)	Kannada
	Kinyarwanda	Lugbara
	Malay	Mandingo
	Mandarin	Nama
	Margi	Nubian
	Maung	Turkish
	Palauan	CARIB
	Sango	DAGA
	Thai	FORE
	Tiv	GUARANI
	Tiwi	HIDATSA
	Vietnamese	KATE
	Yoruba	MARIND
	Zulu	QUECHUA
	OLO	SENTANI
	SULKA	(19)
	YUROK	
	(21)	

(Total languages 50)

Table 4.9d

Languages with post-marking for subject

(non-fusional)

	Pr	Po
VS	-	Bardi
		(1)
SV	-	Bawm
		Burmese
		Burushaski
		Djingili
		Hindi
		Japanese
		Kanuri
		Newari
		Ngandi
		Wati
		GUAYMI
		MAIDU
		NASIOI
		POMO(E)
		SIROI
		WASKIA
		ZOQUE
		(17)

(Total languages 18)

Table 4.10a

Languages with pre-marking for subject (non-fusional)

	NG	GN
VS	Niue (1)	-
SV	-	-

(Total languages: 1)

Table 4.10b

Languages with fusion for subject-marking

	NG	GN
VS	Berber Mâsai (2)	-
SV	Greek (1)	LUISENO (1)

(Total languages 4)

Table 4.10c

Languages with isolation for subject

	NG	GN
VS	Cemuhi	PIRO
	Fijian	(1)
	Maori	
	Welsh	
	Yapese	
	CAYUVAVA	
	JACALTEC	
	SHUSWAP	
	ZAPOTEC	Abkhaz
	(9)	Ewe
		Finnish
		Ijo
		Kannada
SV	Cairene	Lugbara
	Chamorro	Mandarin
	Fulani	Mandingo
	Hausa	Nama
	Iraqw(SOVPr)	Nubian
	Kinyarwanda	Tiwi
	Malay	Turkish
	Margi	CARIB
	Maung	DAGA
	Palauan	FORE
	Sango	GUARANI
	Thai	HIDATSA
	Tiv	KATE
	Vietnamese	MARIND
	Yoruba	QUECHUA
	Zulu	SENTANI
	(16)	SULKA
		YUROK
		(23)

Excluded: OLO (NG~GN)

(Total languages 50)

Table 4.10d

Languages with post-marking of subject
(non-fusional)

	NG	GN
VS	-	Bardi
		(1)
SV	Djingili	Bawm
	Kanuri	Burmese
	(2)	Burushaski
		Hindi
		Japanese
		Newari
		Wati
		GUAYMI
		MAIDU
		NASIOI
		POMO(E)
		SIROI
		WASKIA
		ZOQUE
		(14)

Excluded: Ngandi (NG~GN)

(Total languages 18)

Footnotes

to

Part I

Footnotes to Chapter 1.

1. ND/DN: noun-demonstrative/demonstrative-noun.

2. See Semeonoff (1955).
   In general, grammatical details and/or examples are to be found in the first language reference given in Appendix I for each language.

3. It is not clear whether in fact there is a conflict between my definition of 'semantic apposition' and Quirk et al's account of the semantic content of the appositional construction in English. I understand from Professor Quirk that something along the lines I suggested would be plausible, but would require appropriate examples to back it up.

4. It will become apparent in Chapter 7 (§7.5) that the exclusion or inclusion of 'of' within a region is a matter of indifference as far as semantic considerations are concerned. Further evidence then may well show that 'of' should lie outside the region for English.

   It will be noted that in the ensuing French examples, that the 'de' is treated as lying outside such a region in comparable examples.

5. CNP: Common Noun Phrase.

   That part of the noun phrase (NP) exclusive of the determiner and of post-nominal elements which may modify a singular propername, unaccompanied by a determiner. Thus a restrictive relative clause may be part of the CNP but not an appositive relative clause.

6. Hungarian linguists vary in their description of the pre-verbal expression. Thus Bánhidi et al (1965) term it the emphasised expression, while Kiss (1981) uses the term 'focus'. KIss additionally discusses the intonation pattern.

7. See Tompa (1968).

Footnote to Chapter 2

1. See Hoel (1960), chapter 4: section 2 on the Binomial Distribution.

Footnotes to Chapter 3

1. The examples given here are to be found in sections 721 and 724 of Kroeber and Grace (1960).

2. Examples taken from Li and Thompson (1975).

3. The 'Trigger-Chain' is the term Hawkins uses to describe the hypotheses of Vennemann and Lehmann (see Vennemann (1974).

4. Hawkins actually uses the symbols

   (POST ⊃ (ADJ-N ⊃ GEN-N)).

   However Po, AN, GN, etc are more economic, and more standard to this present work.

5. This chain appears to conform with Canale (1978)'s Implicational Hierarchy Sets (IHS) hypothesis since the ordering of items under a VP node (VO/OV) do not directly correlate with items under an NP node(NG/GN and NA/AN), though there certainly are examples of languages with constructions in which genitive NPs appear to stand outside the NP, eg Hungarian and Pashto. However the method of deriving the hierarchy is different and conflicting results arise on occasion, most notably involving ND/DN (see Chapter 9, § 9.2).

6. The application of Fisher's exact test (see Everitt (1977)) may be illustrated from the left-hand sub-table of Table 3.31:

	NG	GN	Totals
Pr	30	4	34
Po	0	7	7
Totals	30	11	41

I wish to test the hypothesis of a positive correlation between Pr and NG, and between Po and GN.

I must test the probability, given the above row and column totals, that the 41 languages will be distributed as above. This is given by

$$\frac{30! \times 11! \times 34! \times 7!}{30! \times 4! \times 0! \times 7! \times 41!}$$

where $4!$ stands for '4 factorial' and represents $4 \times 3 \times 2 \times 1$, and likewise for $41!$, $34!$, etc. The $0!$ is always interpreted as equal to 1.

The numbers on the top are the row and column totals, and those on the bottom are the cell entries plus the overall total. The above formula is appropriate when one or more off-diagonal cells has a zero in it.

Taking the above formula, we may cancel to get:

$$\frac{\cancel{30!} \times 11! \times 34! \times \cancel{7!}}{\cancel{30!} \times 4! \times 0! \times \cancel{7!} \times 41!}$$

Note that $11!$ over $4!$ gives

$$\frac{11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times \cancel{4} \times \cancel{3} \times \cancel{2} \times \cancel{1}}{\cancel{4} \times \cancel{3} \times \cancel{2} \times \cancel{1}}.$$

We may similarly treat 41! over 34! and the remaining ratio is

$$\frac{11 \times 10 \times 9 \times 8 \times 7 \times 6 \times 5}{41 \times 40 \times 39 \times 38 \times 37 \times 36 \times 35}$$

which may be calculated, either by further cancellations, or by means of a calculating machine, to equal

0.0000146784

or roughly one in 68127.

It is worth looking at another example, from section §3.3. I shall look at Table 3.25, in the left-hand sub-table. This is not in fact a 2x2 table, but it contains the following information for NG languages:

	NA	AN
non-SOV	38	13
SOV	11	0

We have in effect Greenberg and Hawkins' hypothesis that SOV/NG languages are never AN. But is this a significant pattern? The Fisher test gives us a result of 5.73%. One test in 17 or 18 would yield such a result, a perfectly possible result bearing in mind the number of chains I have examined. It must be borne in mind however that the Hawkins sample is not claimed to be geographically or genetically representative.

The formula above may also be used for a negative correlation, if one or more of the diagonal cells is zero. When none of the cells are zero, the formula is a little more complex (see Everitt (1977), pp 15-17), since the test requires that we find 'the probability of the arrangment of frequencies actually obtained, and that of every other arrangement giving as much or more evidence, always keeping in mind that the marginal totals must be regarded as fixed'.

6a. A short note is appropriate on the conflation of Amerindian, Indo-Pacific and 'Old World' languages in the Contingency Tables with identical weightings. To see whether this conceals any significantly different behaviour among languages which might otherwise have a weighting of 2.0, I shall consider the chains

VO/OV ⟷ Pr/Po ⟷ NG/GN

and

Pr/Po ⟷ NG/GN ⟷ NA/AN.

In each analysis, Pr and Po languages are separated. The Pr languages are overwhelmingly 'Old World', whilst a smaller proportion, roughly half to two-thirds of the sampled Po languages, are 'Old World'.

However when separate chain tests are carried out upon 'Old World' and other languages, we find

(a) that 'Old World' languages adhere to the overall pattern,

(b) that the 25 Amerindian and New Guinean languages are only sufficient in number to cast strong doubt on VO/OV and NA/AN as intermediaries.

See Tables 3.5 and 3.26 for relevant data.

The analysis of the chain

Pr/Po ⟷ NG/GN ⟷ VS/SV

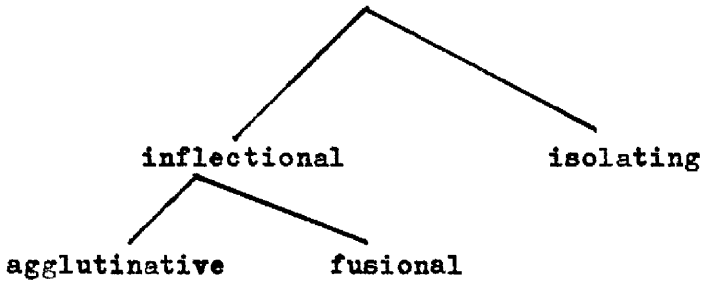
is however much more interesting. Of the 29 'Old World' languages, we find that 17 are Pr/NG/VS, but that none of the six Amerindian or New Guinean languages fall into this group. As a result a different pattern emerges, and it is only apparent among the Amerindian languages that NG/GN is to be preferred over Pr/Po as an intermediary. There is no contrary pattern among the 'Old World' and New Guinean languages: there are simply not enough of the right inconsistent languages to substantiate the pattern: The situation would not be helped if we gave the Amerindian languages a double weighting in some way: we are simply at a point where more data is required.

(See Table 3.30 for relevant data: note that OLO (NG↩GN) is excluded from the above statistics.)

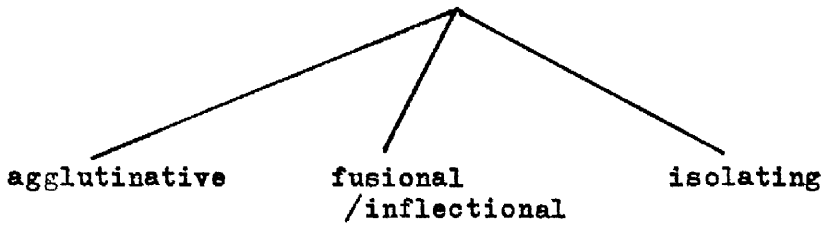
7. Canale (1978)'s Implicational Herarchy Sets (IHS) hypothesis does not make use of a trigger chain: if P implies Q, then as far as Canale is concerned, P does not (necessarily) trigger Q, but rather Q <u>permits</u> P.

Footnotes to chapter 4

1. Along with Comrie (1981b) I use 'inflectional' to subsume 'agglutinative' and 'fusional': thus



rather than



as in earlier work.

However my use of the synthetic/analytic and inflectional/isolating dichotomies differs from that of Comrie(1981b): as far as I can tell, his inflection corresponds to synthesis, with the exception of polysynthesis (see §4.1.1). My own use may be seen from Fig. 1 and its discussion in §4.1.

2. I am indebted to Dr Margaret Bainbridge for checking (and correcting) the made-up examples in the section illustrating the morphology of the Turkish accusative and dative cases.

Note that the 'b' of 'kitabı' becomes 'p' before a consonantal suffix or word-finally. See the second example: kitapları, books(Acc).

3. If it were considered further, we would also have to examine circumfixal constructions, as considered by Greenberg (1980), since there would be no motivation for such fusion if circumfixal constructions were developed.

   It may well be the case that spatial prepositions such as inside, upon, etc are more likely to use circumfixal constructions while cases such as Accusative, Dative, etc will make use of fusion. Such a strategy would at least produce a manageable set of paradigms.

4. See Underhill (1976), pp81-83.

   Note also that Lewis(1967) states that -yle is more common than the Arabic borrowing ve.

   In addition we may note that if case-marking is repeated, NPs may be juxtaposed without a conjunction (see Lewis (1967)). However there appear to be restrictions on how this construction is used (M. Bainbridge, pc.).

5. See Semeonoff (1955) Ex. VIII.B.28.

6. See section entitled 'Contrasting features in /m-/ and minus /m-/ sentences': The m- appears to have similar function to -um- in Tagalog in making the verb agent-focussed. Tung discusses here and in other sections whether the two types of sentences may be treated as active and passive, but concludes that there are insufficient parallels. In particular the agent can never be unspecified in the hypothesised passive sentences.

7. McCormack (1966) also includes in this list yaavdu, which thing? which one? and yaaru, who? after neuter object pronouns and personal pronouns respectively.

   See Unit VII, section 8.

8. See Gregores and Suárez (1968), chapter 15, under 'Object Constructions'.

9. Sources for both GUAYMI (Alphonse (1956)) and NASIOI (Rausch (1912)) use a suffix to mark subject.

Thus for NASIOI:

Utáuna- e bapapa bauran̈ mä́si tampuma
U. her=uncle his=daughter cousin calls
'U. calls her uncle's daughter cousin'

Utáuna- a Barámpanu bamä́ma
U. B. elder=sister
'U. is B's elder sister'

where the -e, sometimes realised as -ke, is treated as a subject suffix by Rausch. This suffix also denoted an instrumental case. No suffix is required for the object. However examples may be found where the subject of the intransitive no suffix:

tavi ntoṅ ke nä́mpearama
fish water in swim
'Fish swim in (the) water'

(an example used to illustrate the postposition ko).

Similarly in GUAYMI we find the suffix used for animate subjects in the

ti etebä- we ñebare
my brother said=it

But we also find subjects of intransitive verbs without the suffix -we:

9. (cont'd)

kurá dobún gotobitibare slota bori baliente
tiger fierce met boy very brave

ben kodrunente
with in=the=dark

'A very fierce tiger met with a very brave boy in the dark'

(As it happens, the example is used to illustrate the comparison of adjectives: bori, very, more.)

For neither language does the source explain the absence of the 'nominative' suffix in these examples and subsequently the possibility of ergativity in these languages must be considered. The evidence in NASIOI is not incompatible with that of WASKIA, where the Ergative marking ke may also be used for human subjects.

10. McLendon (1975) states that intransitive verbs require no morphological marking. In a later discussion however (McLendon (1978)), she exemplifies 'intransitive' verbs of falling, being burned, forgetting, etc. which take object marking:

mi'pal p^{h}a'béka 'He got burnt'
him got=burnt

But this does not affect the tripartite nature of EASTERN POMO, since there remains a body of intransitive verbs, such as wá-du'kì, go away, which take neither ergative nor object marking.

11. See footnote 9, also concerning NASIOI.

12. I am grateful to Mr Pratap Singh for checking (and correcting) these two made-up examples, which I hope succinctly illustrate the point made in the text.

Examples are also to be found in McGregor (1977), illustrating both perfective Tenses and (-)ne, and include the following:

usne	un	striyoṁ	ko	pahle	dekhâ	thâ
he	those	women	Obj	before	seen	was
(Erg)	(Obl)	(Obl)				

'He had seen those women before'

maiṁ	kal	âyâ	thâ
I	yesterday	come	was
(Nom)			

'I came yesterday'

Although ne is normally a postposition, following an NP in the Oblique Case, it combines fusionally with pronouns, such as vuh, he/she/it/that, in the first example.

13. I am grateful to Mr Hailu Habtu for this example. But note that according to Praetorius (1871), the 3sg masculine object suffix is much more commonly -ô than -u, and it is possible either that I misheard the vowel, or that Tigrinya has changed in this respect over the last 115 years or so.

14. I am grateful to David Hawkins and Susan Rollin for these examples.

15. But see Ergunvanlı (1984), chapter 2. It appears here that under certain circumstances one part of a larger NP may be placed after

15. (cont'd)

the verb.

16. The two exceptions to the case-agreement implication, Berber and Iraqw, do not show any inconsistency with verb position: thus for the VS language Berber, the case-marking is prefixal, and for the OV language Iraqw, object-marking on the noun is suffixal. The implication might therefore be refined, if more data justified it, so that only fusion motivated by inconsistent ordering implied case-agreement.

PART II

SEMANTIC CONSTRAINTS

5. Syntactic universals and semantic constraints

In the last two chapters I have discussed the behaviour of syntactic patterns: chapter 3 in effect looked at patterns involving two or more lexical items, while chapter 4 concentrated more on the interaction of lexical and grammatical items. I now wish to look more closely at semantic constraints, and it is therefore time to examine in greater detail the work already done on semantic structures, especially the intensional logic (IL) of Montague, since Bartsch and Vennemann regard their operator-operand approach of Natural Serialisation as a development of Montague's work.

The general thrust of attempts to explain word- and morpheme-order patterns is to devise a concept of 'consistency' whereby 'modifiers' consistently precede or follow their 'heads'. It will be my task in these coming chapters, 5 to 8, to show that for both syntactic and semantic reasons, more than one type of 'modifier' must be identified: the most typical of the first type of modifier is the pre- or post-positional object, especially when joined to its head pre- or post-position by simple juxtaposition, the most typical of the second type is adjectives, especially when they may freely precede or follow their noun, and the third form, where the concept of modifier and head is more problematic, is typified by deixis, in the form of determiners and perhaps in the form of tense.

The main task of this present chapter will be to discuss the work of Montague, relevant to the present study, with a brief mention of Bartsch/Vennemann in my discussion of the problems posed in trying to use IL or something like it for explaining word-order patterns. Against this background I shall further look at semantic considerations in chapters 6 to 8.

Chapter 6 is a very lengthy chapter on the semantics of adjectives and indeed other nominal adjuncts: it will emerge from this discussion that there is indeed a correlation between ordering patterns of adjectives and their semantics, and that we can <u>partly</u> account for this by intensional logic. Arguments are then introduced for the two types of modifier, whose properties are then discussed in chapters 7 and 8, the latter extending the discussion of adjuncts from nominal to verbal adjuncts.

For the present purposes we may regard IL as a device whereby similar

syntactic structures may be translated into very different logical structures: a sentence such as

John eats doughnuts

lends itself readily to predicate calculus, while

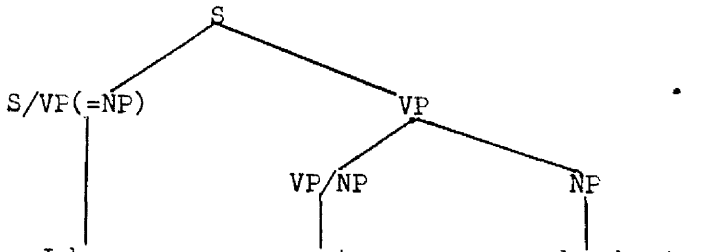
John hunts unicorns

does not:

$$(\exists x)\ (Px \ \& \ Qjx)$$

could x stand for the former proposition - with j standing for 'John', P for 'is a doughnut' and Q for 'eats' - stating that there are x's which are doughnuts and John eats them: the latter sentence cannot be represented in this way since it does not mean that there are unicorns and John hunts them.

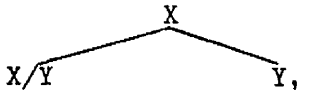
It is this question that Montague set about in his paper 'The Proper Treatment of Quantifiers of Ordinary English' (1970b) - also referred to as PTQ - and within his framework the above sentences may take the following structure:



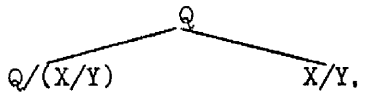
We may note that the syntactic categories follow a 'categorial grammar' (see Dowty (1975)): an X/Y is something which combines with a Y to form an X.

The above tree shows us that a transitive verb (V_{tr}) is treated as VP/NP: it requires an NP to form a VP. We may also note that a noun phrase appears either as NP or S/VP. The S/VP is in fact the basic one, and NP defined in terms of it, in effect as a potential subject. This means that an item X/Y may be introduced in two ways:

(a) under an X node as a sister of Y:

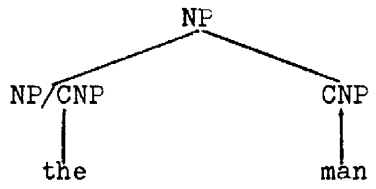


or (b) as a sister of something with (X/Y) after the oblique:



It is a critical characteristic of the Montague approach that an NP is primarily defined in terms of its role as a potential subject. A transitive verb is defined then as an item requiring a potential subject (S/VP) to make up a VP. This brings up the question of which item in a two-element phrase becomes X/Y and which simply Y, and I shall discuss this when I consider the semantic interpretation of the above structures (later in this chapter), but we may note at this point the advantage of treating the NP as X/Y in one environment only, that of subject, rather than multiply defining it as VP/V, PP/P, etc.

We may now consider other categories. If we identify a CNP as the NP apart from the determiner, Det may be treated as NP/CNP:



In the present study I have brought the notation into line with the usual PSG conventions. We may thus compare

PTQ	PSG
CN	N or CNP
IV	V_{itr} or VP
T	NP or propername
TV	V_{tr}
IAV	Adv or AdvP or PP

Thus in any discussion of Montague categories, VP corresponds to IV (intransitive verb) in PTQ, CNP to CN and AdvP to IAV. This is because rules in PTQ applying to VP's apply equally to intransitive verbs, etc. (The IAV stands for 'intransitive adverb', so that a 'transitive adverb' would in fact be a preposition. T stands for 'term'.)

In addition to the above categories we may identify an important group, that is one where the items before and after the oblique are identical, X/Y:

CNP/CNP	adjective
VP/VP	verbal adverb
(CNP/CNP)/(CNP/CNP)	adjectival adverb

Sometimes more than one slash is necessary: VP/VP stands for a verbal adverb, while VP//VP stands for a verb taking an infinitival VP rather than an NP. It is entirely arbitrary which is chosen as VP/VP and which is chosen as VP//VP.

The question arises of how these combine semantically: in fact we generally have a very simple rule, that the meaning of X/Y represents the regular relationship between the meaning of Y and the meaning of X. Put more formally, the intension of X/Y is the function which maps the intension of Y into the intension of X. Let us look more closely at this concept of intension, and in particular the distinction between intension and extension, which corresponds to Frege's distinction between Sinn and Bedeutung. The intension of an expression is its basic meaning, while the extension is that to which it refers at a particular index. The meanings of the expressions 'mother of Charles' and 'Queen of England' may be different, but because 'mother of Charles' and 'Queen of England' refer to the same person, we may equally well say that the Queen of England watched the races, and that the mother of Charles watched the races. But we cannot equally well say:

The Queen of England may sack the Prime Minister

and

The mother of Charles may sack the Prime Minister.

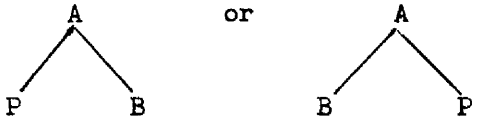
What we say here is that the truth of the first pair (watching the races) essentially only depends upon the extension, though it could also be reconstructed from the intension, of the subject NP of these expressions, whereas in the second pair (sacking the PM) the truth values are essentially dependent upon the intension, and we cannot reconstruct the intension from the extension. Note that it is the word 'may' that creates the problem: if the Queen of England actually did sack the PM, then so did the mother of Charles.

It would appear that most X/Y categories include intensional members, as may be seen from the list in Dowty (1975):

PTQ category	Description	Examples
S/S	sentential adverb	Necessarily...
VP//VP	verbs taking infinitive, incl. modals	must
VP/S	verbs with sentential complement	believe
VP/NP	transitive verbs	seek
CNP/CNP	adjectives	alleged
VP/VP	verbal adverbs	allegedly
AdvP/NP	prepositions	about

These intensional adjectives, verbs, adverbs, etc. contrast with extensional counterparts, such as did, find, white, in, etc.

However, these extensional words may also be taken to depend upon the intension. For this reason Montague chose to build up the meaning of phrases in terms of their intensions. How do we express this? For phrases such as those discussed above, which may be broken down into



where P is A/B or A//B, we will have a translation rule giving us the following semantic expression:

$$\underline{A}' = \underline{P}'(\hat{}\underline{B}')$$

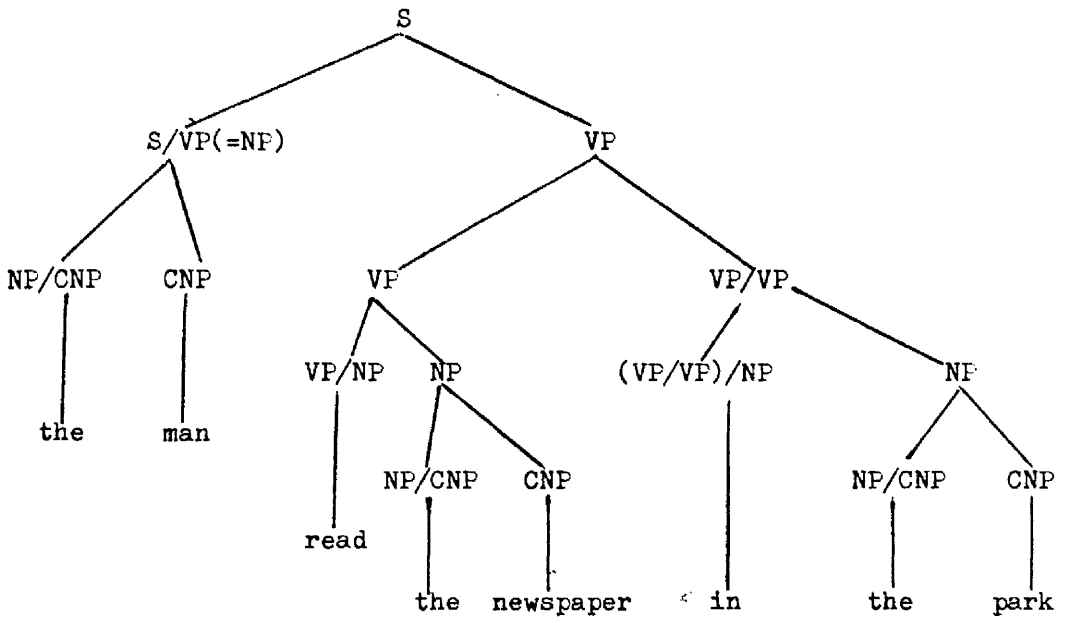
where $\underline{B}'$ is the intension of $\underline{B}'$ and in general $\hat{}\underline{x}$ is the intension of $\underline{x}$. Thus the $\underline{P}'$ is a function, and the bracketed expression its argument.

By way of illustration we may consider the sentence:

The man read the newspaper in the park,

which comes out as

(see next page)



We may start by analysing the subject NP: by invoking the principle

$$\underline{A}' = \underline{P}'(\hat{}\underline{B}')$$

where P=A/B (or A//B), we get

$$\underline{NP}' = \underline{the}'(\hat{}\underline{man}').$$

We may now analyse the complete sentence (S), which comes out as

$$\underline{S}' = \underline{NP}'(\hat{}\underline{VP}')$$

$$= \underline{the}'(\hat{}\underline{man}')(\hat{}\underline{VP}').$$

We may now carry on and analyse the VP expression 'read the newspaper in the park', and here we obtain

$$\underline{VP}' = \underline{(VP/VP)}'(\hat{}VP').$$

Here the VP/VP stands for 'in the park' and the second VP stands for 'read the newspaper', which may now be analysed in the same way as NP and VP under the S node, to give the full result as follows:

$$\underline{the}'(\hat{}\underline{man}')(\hat{}\underline{in}'(\hat{}\underline{the}'(\hat{}\underline{park}'))(\hat{}\underline{read}'(\hat{}\underline{the}'(\hat{}\underline{newspaper}')))).$$

Corresponding to each translation rule is a syntactic rule. As an example, the syntactic rule combining a third person singular subject NP with its VP specifies firstly that the NP precedes the VP, and secondly that -s is added to the main verb. It must be emphasised however that Montague's work in the area of verbal inflection and its associated syntactic and translation rules was ad hoc and highly tentative (see Dowty (1981) and his concluding remarks in chapter 7,

on the grammar of PTQ).

Above is an account of PTQ's most important principle of phrasal interpretation. The importance of Montague's work to linguistics has undoubtedly been diluted on account of its highly complex presentation - and some people say its appalling syntax - and the above account has attempted to elicit the basic principles and issues that are relevant to the present work. Although PTQ is language-specific in the sense that English is the language under discussion, the principle of translation into intensional logic may be and has been applied to other languages. The principle that adjectives are of category CNP/CNP is not a language universal however, as Bartsch and Vennemann appear to assume: it must be justified in each and every language that has adjectives, in accordance with the way that adjectives are used in that language. It must be further noted that when Montague is concerned with universal grammar, however, he is concerned with a mathematical concept of translation, which applies to any language, and here intensional logic as such is simply one of the things that may be captured within language.

It will be seen that Montague Grammar does not as it stands provide an explanation for syntactic universals as observed by writers such as Greenberg and Vennemann. Not merely does IL have to be valid as a theory of logic, but it also has to be shown to be relevant, and here it is necessary to provide statistical evidence for the phenomena to be explained. An example of a situation where the validity of IL is not questioned but the relevance is, comes in J. Kamp's article on 'Two Theories about Ajectives' (1975), where an alternative theory of adjectives is put forward (see ch. 6), in which adjectives are treated as if they are all extensional, with contextual factors taken formally into account. This does not prevent an intensional approach being used in other constructions.

This work will then assume the utility of IL, and of the Montague programme in general (in particular its system of categories). To do otherwise would be problematic when we are primarily concerned with his principle of phrasal interpretation. What must be done however is to examine the evidence from a representative cross-section of natural language, and then examine the degree to which it is relevant and economic in accounting for the statistical universal patterns found in such investigations.

In turning to the original phrase structure interpretations, it might at first be tempting to argue that here we have a motivation for word-order universals: in a VO and Pr language we get

burn'(^the'(^cake')),
in'(^the'(^park')).

If a language has VO and Pr, with structures such as the above, surely this means that in 'VO languages', functions precede arguments, whereas in OV languages they follow? In effect, when an expression A can be broken down into A/B and B, the A/B must precede in VO language, but follow in OV language.

If we are to take this approach and relate it to the statistical evidence of the last two chapters, we must in effect show how it relates to the prediction chain. We would want something to the effect that if a prediction chain has the form:

VO/OV $\longleftrightarrow$ Pr/Po $\longleftrightarrow$ NG/GN $\longleftrightarrow$ NA/AN,

languages will tend to be consistent with adjacent items on the chain in placing function before argument or vice-versa. There are however a number of difficulties with this approach, and three in particular may be cited:

(i) Adjectives, as well as prepositions and transitive verbs, are functions, predicting that AN is consistent with Pr and NG (as with English), and NA with Po and GN. In fact the reverse is true, at any rate for Pr and NG, with NA.

(ii) PTQ treats adverbs as functions, yielding intensional logic structures such as the one illustrated above:

the'(^man') (^in'(^the'(^park')) (^read'(^the'(^newspaper'))))

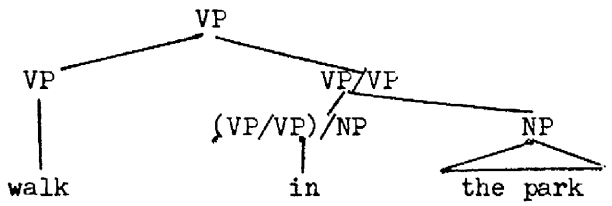
for 'the man reads the newspaper in the park'.

In effect we predict that VP/VP adverbial expressions will precede the VP in a VO language, and follow it in an OV language. Again we have a wrong prediction.

(iii) Correlations are very uneven when they occur: while the correlations for VSOPr, SOVPo, Pr/NG, and Po/GN are very high, the correlations between adjectival and genitival ordering are of a different nature, with NG languages strongly NA, and GN languages evenly split. The same thing may also be observed with VO/OV and Pr/Po: OV languages are strongly Po, while there is a more

even distribution of VO languages between Pr and Po.

But if we look more closely at the offending expressions - adjectival, adverbial and prepositional phrases - it will be noted that in each case we have not so much an A/B expression as an expandable X/X expression:



(where we may again note that VPs and intransitive verbs are treated identically).

This particular problem appears to have motivated the Bartsch/Vennemann approach where the object acts as operator upon its verb to form a verb phrase (VP/V), and indeed all modifying elements except perhaps the subject NP are operators: all in effect expandable X/X or XP/XP (eg. CNP/CNP, VP/VP), or perhaps XP/X in the case of NPs: VP/V, PP/P, etc.

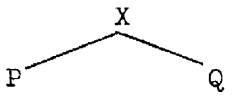
Difficulties emerge with both the Montague and the Bartsch/Vennemann approaches, and in each case it involves the correlations observable in X/X expressions with intensionality. In the case of Bartsch/Vennemann the problem, as I outline in §6.3, is straightforward: those adjuncts which most require an intensional analysis, ie. those which most require to be treated as operators or functions are those which are most likely to violate Natural Serialisation. In the case of the Montague approach not only do we have difficulties with expandable X/X expressions as discussed above: we must also recognise that if correlations occur with intensionality, there must be constructions where an intensional approach is not appropriate. In effect, even though an intensional analysis may be extended to 'extensional' semantic structures, we should not take advantage of this to conflate syntactically similar structures under an intensional analysis.

However, because of the 'fail-safe' nature of IL - its ability to reconstruct an extensional analysis via an intensional analysis - we are not refuting IL simply by questioning its appropriateness for certain constructions: instead, by focussing on a limited number of such forms, the process is similar to that of peeling off one or two skins

from an onion. This is as well in fact, since it spares the present work of the danger of having to say something about all constructions in natural language.

Turning then to adjuncts, we are fully entitled to ask whether a separate strategy is appropriate for their explanation. Is it because of something special about the semantics of X/X? Or is it because of its expandability? Or are the two connected?

It would at first sight appear to be the case that if expandability is the key factor, word-order universals are in reality syntactically motivated, so that if we have the expression



where Q is expandable but P is not, eg. P is a preposition and Q a NP, questions of ease of parsing will in all likelihood arise. If on the other hand, it is something about the nature of X/X itself, then we must look more closely at the alternative ways of semantically interpreting X/X expressions. If, however, the two are interrelated, we might at first sight think that we are unable to detach semantic and syntactic motivations. It will be my contention that the two are indeed interrelated, but that the situation is not indeterminate, but rather that separate influences are discernable in different types of inconsistency.

6. Adjectives

0.

I shall now turn to the question of semantic constraints with specific reference to adjectives. To do this, I shall examine the evidence from my language sample, as to the ways in which the semantic characteristics of adjectives influence their syntactic behaviour. In chapters 3 and 4 the following syntactic points were made about adjectives:

(i) adjectival ordering correlated much more strongly with NG/GN than with VO/OV, and in Hawkins' sample more strongly with NG/GN than Pr/Po,

(ii) this correlation was rather weaker than that observed between NG/GN and Pr/Po, or between Pr/Po and VO/OV,

(iii) when case-markings became fused, case agreement appeared to be necessary to avoid fusion of attributive adjective and noun.

In this chapter I wish to put forward evidence to show that in semantic terms, ordering of adjectives correlates most strongly with pre-/postpositionality, in that when a language is NA, certain types of adjectives, usually dimensional, may still precede the noun, while among postpositional languages, it is another set of adjectives, including nationality, that are most likely to precede.

It is important to note the difference here: on a syntactic level we have a prediction chain where NA/AN correlates with NG/GN, though my last chapter will show that this occurs indirectly via ND/DN. But when we look at the semantic properties, correlation appears to occur instead with Pr/Po, certainly among descriptive adjectives. The evidence for this assertion will be presented in the form of an adjectival hierarchy, at the top of which appear items most likely to precede in Pr languages and follow in Po languages, and at the bottom items most likely to follow in Pr languages and precede in Po languages.

It is not the aim of this present study to present conclusive

evidence for the ordering of items on the hierarchy. A sample of 75 languages is sadly not sufficient to do this successfully. The most important task in this present work in using the hypothesised hierarchy is to show that if one wishes to make use of a Principle of Natural Serialisation, such as that of Bartsch and Vennemann, in which intensional adjectives are treated as operators, it is not sufficient to assume that anything following in an ordering consistent with NA will also be an operator. The objective of the adjectival hierarchy used here is to show that it is necessary not merely to consider adjectives as a whole, but to investigate the variation in ordering among different semantic types of adjectives. As will be shown in the next chapter this further investigation will suggest that for many constructions the identification of operator and operand is the opposite to that proposed by Bartsch and Vennemann.

A small note should be added about what constitutes an adjective in classifying a language as NA or AN: while it is accepted by Greenberg that the criterion must be semantic, little is said about the precise criterion, which on occasion must be crucial in a work such as this. To capture the present practice, which for example treats Hausa as AN and Burmese as NA (see below), I treat adjectives as the class including the descriptive 'big', except where this can be identified as an Augmentative/Diminutive class (eg. Cemuhi, Tiv), in which case the adjectival class shall include 'good'.

6.1 Types of adjectives

So far in this work adjectives have only been distinguished at what might well be an extra-linguistic level: how the noun (if any) that they modify is to be taken into account when an appropriate logical form is to be considered.

Thus in the phrase 'black sheep' we only need to consider the extension of black objects and the extension of sheep in order to obtain the extension of black sheep. In the phrase 'fake pound note' however, we cannot rely on the extension of faked objects and of pound notes: a fake pound note is only a fake with respect to its status as a pound note - it is a genuine piece of paper! It is however debatable whether the extension of pound notes includes fakes, but in fact the problem does not arise in the Montague approach, since by making 'fake' a function of the intension of 'pound note', this is sufficient to determine both the extension and intension of 'fake pound note'.

Does this dichotomy at a possibly extra-linguistic level correspond to any similar dichotomy at a level that is more clearly linguistic? It would appear not, if the intensional analysis is to cover everything. However Siegel has attempted in her work 'Capturing the Adjective' (1980) to discuss syntactic differences in English, Russian, Italian and Ngamambo which might be motivated by an intensional versus extensional distinction. Such differences include those in English, where she finds that all 'extensional' adjectives may be predicative, whilst all 'intensional' adjectives may be pre-pre-nominal.[1]

The intensional adjectives, or adjectives used with their intensional meaning, would have a basic pre-nominal form, with the predicative form derivable from it. Siegel puts forward the rule

$$\text{be a} \quad [\text{ADJ}]_{\text{CN/CN}} \quad \triangle_{\text{CN}} \longrightarrow \text{be} \quad [\text{ADJ}]_{\text{CN/CN}}$$

where the item $\triangle$ is a dummy common-noun, which may alternatively 'come to the surface' as <u>one</u> for singular count nouns. Thus the sentence

This diversion is temporary,

is treated as being similar in meaning to

This diversion is a temporary one.

Conversely extensional adjectives may be derived from predicates so that 'an angry man' has the same semantic interpretation as 'a man who is angry'. This is very similar to the device preferred by transformationalists, of 'WHIZ-deletion'.

Siegel notes, however, that 'measure' adjectives seem to have properties similar to extensional adjectives: thus 'Little Dorritt' is more like 'Dorritt who is little' than somebody who is 'little for a Dorritt'. If this is the case, her approach is, as she observes, entirely consistent with Kamp's second theory of adjectives in his (1975) work on 'Two Theories about Adjectives': the examples he considers are in fact 'measure' adjectives, and they are treated as extensional with allowance made for context.

This chapter will be concerned with the evidence available from my sample of 75 languages, and unfortunately data on questions such as the acceptability of sentences resembling, say,

This road is main,

is very difficult to get hold of.

Consequently I shall now consider a less theoretical and more overt classification of adjectives, which throughout this chapter I shall consider as a hierarchy in terms of likelihood to precede (AN) in prepositional languages, and likelihood to follow (NA) in postpositional languages:

Group	Description	Examples
A	Indefinite	some, all, such, another
B	Quantitative	few, many, three
C	Ordinal (incl. Superlative)	earliest, third
D (i)	Modal	possible, former
(ii)	Age	old
E	Dimension	small, long
F	Value	bad
G	Speed	fast
H	Colour	red
I	Physical Property	round, hard
J	Human Propensity	angry

Group	Description	Examples
K	Material	wooden
L	Nationality	Abkhaz

We may note that Siegel's 'measure' adjectives appear to span E to G. Thus Siegel discusses tall (E), good (F) and fast (G) in her chapter on measure adjectives and their quasi-extensional behaviour.

As mentioned above, being in group A means that an adjective is more likely to precede than an adjective which is not in group A. Thus we shall see at a later stage that items denoting all precede in approximately 72% of prepositional languages, while numerals (group B) precede in approximately 68% of Pr languages, while items denoting small (group E) precede in approximately 36% of Pr languages.

The fact that numerals (B) appear to the top of the hypothesised hierarchy means then that if a language is postpositional, but places numerals before the noun, virtually all the other noun modifiers listed will also precede. In fact virtually all Po/NumN languages in the sample are also AN, placing 'descriptive adjectives' before the noun, the exceptions being GUARANI (NA), LUISENO and PIRO (NA~AN). Similarly we may note that among Pr languages, all but two of the NNum languages, ie, those that placed the numeral after the noun, were also NA. (Exceptions Hausa and Sango.)

Similarly if Dimension adjectives (E) follow in a postpositional language, items higher up will also tend to: in fact all such languages in this sample (except GUARANI) place numerals after the noun (NNum). Conversely if Dimension adjectives precede in a prepositional language, items higher up will also tend to precede: if we again compare this order with numeral position, we find this to be the case, with one clear exception, Tiv, plus Hausa, which normally places such adjectives first.

This hierarchy parallels the Prepositional Noun Modifier Hierarchy of Hawkins (1979), which posits a hierarchy of Demonstratives, Descriptive Adjectives, Genitive NPs and Relative Clauses. My hierarchy differs in three respects:

(a) the only overlap of items occurs with descriptive adjectives (though my suspicion is that relative clauses could also appear on my hierarchy, at the bottom).

(b) no attempt is made to produce an exceptionless hierarchy, as exceptions cited above illustrate,

(c) the hypothesised hierarchy applies to postpositional languages as well as prepositional (although the results are rather more blurred among Po languages).

It might be thought from the above discussion that we may formulate implications of the form:

$$Pr \supset (NA_x \dot{\supset} NA_y),$$

and $$Po \supset (A_xN \dot{\supset} A_yN),$$

where NA_x is above NA_y in the hierarchy.

This does not in fact follow at all. Thus I observed that my sample suggested that around 72% of prepositional languages prepose 'all', whilst 68% of Pr languages prepose numerals. But it is only when there is a very high correlation between the placing of 'all' and the numerals that we can depend upon an implication of the above form. In this chapter however, only the ordering on the hierarchy is critical, and even here only the broad ordering. The question of whether implicational rules can or need be derived will be deferred to the last chapter, where I shall consider more closely the question of adjectives and their position on the prediction chain.

It will emerge from the discussion of this chapter that the ordering is broadly borne out. It must be pointed out however that the pattern is rarely so precise that the relative ordering of any two adjacent items on the hierarchy can be confidently predicted. Thus the ordering of A and B is uncertain, as are C and D, I and J and K and L. In addition, in one particular instance, that of G (Speed), the position of it is even less certain: it could well come after J.

The hierarchy can only be discerned from languages which place some noun modifiers before and others after: the number that do this among postpositional languages in the sample is only a third of those with varying orders among prepositional languages. As a result, it is really only A and B at the top, and K and L. at the bottom, that stick out at all for postpositional languages. To verify the finer detail of the hierarchy for postpositional languages would thus require a trebling of the sample.

The above classification of adjectival concepts into the above twelve

groups is essentially for the purposes of identifying patterns rather than identifying universal adjectival categories. It draws heavily from the work of Dixon (1977), who examines pre-head elements in English and considers them to adhere to the following unmarked order:

I. pre-adjectival modifiers

logical qualifiers	(all, some, etc.)
determiners	(the, this)
possessives	(my, John's)
superlatives	(best, cleverest)
ordinals	(fourth)
cardinal numbers	(four)

and so on.

II. adjectives

1.	Value	(cf. F)
2.	Dimension	(cf. E)
3.	Physical Property	(cf. I)
4.	Speed	(cf. G)
5.	Human Propensity	(cf. J)
6.	Age	(cf. D(ii))
7.	Colour	(cf. H)

III. post-adjectival modifiers

origin/composition - eg. oatmeal dog food

purpose/beneficiary - eg. oatmeal dog food

In my own analysis I am therefore widening the concept of adjective. I am doing this for two reasons: firstly there are many languages which treat other concepts, such as logical qualifiers in a similar syntactic manner to 'descriptive adjectives'. Secondly the wider classification is relevant to attempts to provide a semantic motivation for the syntactic behaviour of adjectives, whatever these may include: adjectives of nationality will be particularly relevant to the discussion.

In particular I am breaking down Dixon's first group of 'pre-adjectival modifiers' into A, B and C. (This concept is of course debatable: many linguists would regard superlatives and ordinals as adjectives and the rest determiners. However it may very well be true, as he implies, that superlatives and ordinals normally precede other adjectives.) I have two clearly defined groups, B and C, and a rather vaguer one, A, which does not however include articles, demonstratives

or possessives, which will be considered separately.

B primarily includes cardinal numbers, but I am also including plural quantitative items such as few, many, etc. I have lumped the latter into this group since it seemed highly probable that they would behave in a similar syntactic and semantic fashion, and that if this were not the case the fact should be spotlighted. Part of my task has been to classify my 75 languages as NumN or NNum, depending upon the position of such items before or after the noun, and to do this most effectively I have excluded 'logical qualifiers' such as some, all, most from B.

C includes both ordinals and superlatives. It is certainly not the case that these are identifiable as attributives in all languages, certainly not in my present sample, but there appear to be similarities of treatment in some languages. It is possible that ordinals should come higher in the hierarchy than superlatives.

This leaves the rather vague group A, in which I am including the logical qualifiers. I have essentially picked out what seem to be grammatical items discussed in the grammars, often under the title 'Indefinite Adjectives', which in the following pages include:

I	II
all	also
any	different
another	alone
a certain	himself, etc. (Intensives)
each	separately
same	similar
such	
various	

A number of the above would not be treated as adjectives in English, but they seem to occur as such in certain languages - in the same way as English treats certain concepts as adjectives though some other languages would not do so.

The miscellaneous nature of the grouping means that there are bound to be exceptions in it to my hypothesised hierarchy, and this may be seen in the fact that the items selected for the right-hand column follow their noun in French (aussi, divers/différent, seul, même, séparément, pareil), even though there are plenty of preceding adjectives in French among those classified lower in the hierarchy, especially B to F. A

detailed analysis of type A 'adjectives' would in fact be a comprehensive task, and I basically allow exceptions to ordering patterns under A to emerge and comment upon them towards the end of this chapter.

Groups D to J are simply taken from Dixon's classification of adjectives, with the addition of D(i). This group is named Modal, but strictly includes temporal concepts as well, as will be seen in the following examples:

false beard
real beard

possible solution

obvious trick

old soldier

former PM
present PM

future PM

They are concerned with the question of whether the object referred to really can be described by the description of the head noun. An angry white rabbit is both angry and white, but a present GLC councillor who is also a future PM cannot likewise be described as both present and future, unless we infer something additional, such as 'big name'.

Such adjectives are therefore strong candidates for intensional treatment, as indeed are Groups E to G, though for different reasons: in talking about a small elephant we are not concerned about whether the elephant is an elephant, but the concept 'small' is still relevant to the noun used. Considering the question in intensional/extensional terms, we may pick out the same object with the description 'elephant' or 'creature', but if we describe an elephant as a 'creature' we can no longer describe it as 'small'. Similarly for someone who is a good clergyman or carpenter, but a bad darts-player.

The placing of speed, G, together with E and F must be remarked upon in that in phrases such as 'fast tortoise' we can certainly see the intensional use of 'fast'. On the other hand, we get expressions such as 'fast runner': admittedly a slow runner may be faster than a fast walker, but it is perfectly feasible that a quasi-adverbial grouping

would be more appropriate, where 'fast runner' means someone who runs fast, and 'beautiful singer' means one who sings beautifully, regardless of his/her good looks or lack of them. (In many languages in fact, according to Dixon, speed expressions can only be adverbial.) Such an approach would however be extremely difficult to test, trying for example to distinguish between 'ordinary' and quasi-adverbial Value adjectives. Consequently I have simply grouped Speed adjectives together with Value and Dimensional adjectives on the hypothesised hierarchy. As observed above, however, the results from the data on group G are marginally negative for this approach, and it is difficult to draw any conclusions for this group without a wider study of adverbial concepts than appears in this present work.

Groups D(ii), H, I and J are classes which are more eligible for both extensional and intersectional analysis: extensional because an angry white rabbit is angry and white whether it is described as a rabbit, creature, thing or whatever, and intersectional since the set of angry white rabbits consists of the intersection of three sets, that of rabbits, that of white objects, and that of angry 'objects'. None of the adjectival groups earlier described lend themselves either to extensional or to intersectional treatment, unless some contextual strategy is devised.

The reason for placing D(ii) above E, F and G in the hierarchy is that 'old' is often either synonymous with 'former', or the references list a word as meaning 'old' without distinguishing its various meanings. If the reason is that homonymy occurs in a large number of these languages, then the concept 'old' is simply pulled up the hierarchy by its other meaning. It must be noted however that the placing of D(ii), Age, so high in the hierarchy is based solely on the evidence of 'old': the opposite expression 'new' and 'young' could well occupy a lower position.

The remaining groups of adjectives on the hierarchy are materials (K) and nationality (L). They correspond approximately in meaning to Dixon's origin/composition post-adjectival modifiers, but it has been necessary to consider them closely, since they have certainly appeared among minority AN adjectives in NA languages. I have suggested that they are most likely to follow in Pr languages and to precede in Po languages. It may well be the case that the critical factor is rather the genitive position, NG/GN: many languages use genitival constructions for one or both of these groupings.

Though I have only tested Nationality in this work for group L, it may be that we should be looking at a wider concept, perhaps termed Generic or Referential, including religion, political affiliations, styles, fashions, etc., eg. 'the Welsh leader', 'the Labour leader', 'the Beatle craze' and so on.[2] The examples suggest that such a Generic/Referential concept merges with proper-names, but maybe words such as 'strange', 'heterogeneous', 'similar' and 'novel' should also be included. If we take nationality alone, we find a motivation for specific syntactic or lexical items to adopt a particular ordering, as in Basque, where the word for 'Basque' is the only 'adjective' to precede, while Abkhaz and Waskia place any adjective of nationality before the noun. However the wider concept would explain why certain adjectives including 'indefinite' follow the noun for certain meanings in French: consider pareil, divers, nouveau meaning 'such (A?), 'various' (A?), 'new' (D(ii)) before the noun, but 'similar', 'different' and 'novel' after the noun.

Examining the adjectival groups among 75 languages is a somewhat daunting task, and my strategy has taken the following form:

(a) to establish the dominant ordering (if any) for the position of attributive adjectives (ie. NA or AN) where the adjectival category is defined as in the introduction to this chapter, observing deviating adjectives where they occur,

(b) to establish the construction for the quantifier 'all', and its dominant ordering with respect to the head noun if it attached to the relevant NP (allN/Nall),

(c) to establish the construction for cardinal numerals and their dominant ordering with respect to the head noun if they are attached to the relevant NP (NumN/NNum).

(d) to establish the position of the attributive adjective (or equivalent) meaning 'small' (smN/Nsm).

The (a) to (d) above represent the attempt in this survey at an exhaustive analysis of adjectival position. Other details, such as the behaviour of semantic equivalents of adjectives are discussed when data has been encountered and found to be relevant. Although only 'small' is singled out as a descriptive adjective, the behaviour of other descriptive adjectives, under D(ii) to F at any rate, also emerges from the discussion.

So far I have considered potential adjectival concepts in terms of their inherent semantic characteristics. For the purposes of word-order, however, other criteria must be borne in mind:

(a) Lexical: in a number of languages, the position of the adjectives is simply part of the information that goes with the lexical item - some follow, others precede and still others follow or precede with a change in meaning. This last pattern clearly emerges in the Romance languages, where in French, with some adjectives, it appears that variation in position goes with a difference in meaning, eg. bon, meaning 'good' before the noun, 'le bon voleur', but 'kind' when following: 'le voleur bon'.

(b) Syntactic: more complex adjectival expressions may in some languages have a greater tendency to follow than simpler or one-word expressions. Thus according to Ferrar (1967), a French adjective may precede the noun when modified by an adverb, only if that adverb is très, bien, fort, plus, moins, assez, aussi or si. It may well be the case that the tendency for relative clauses to follow in prepositional languages and precede in postpositional languages follows such syntactic considerations.

(c) Pragmatic: under this heading we may consider emphasis and contrast. If we take the phrase 'big box', either noun or adjective may bear emphasis, and one purpose of such emphasis may be to form a contrast with, say, a small box, or a big pot. In Spanish, for example, the adjective follows if it is the contrastive item, but precedes if the noun is.

(d) Semantic (non-inherent): here we may consider the distinction between restrictive and non-restrictive or appositive. This distinction is important in English, since restrictive and non-restrictive relative clauses differ in their syntactic behaviour (at least in definite NPs).[3]

It is thus clear that the criteria governing the adjectival position is not due solely to the semantics of the adjective itself. In particular, most languages where a dominant order is identifiable among alternatives used within the language have lexical criteria for the place of each adjective and it is only by cross-language comparison that the semantic pattern emerges. It is because of the scattering of odd lexical items through the grouping on the above hierarchy that we

cannot expect the rigid hierarchic or implicational principle where for example if group C adjectives precede so must B.

In English we find lexical, semantic and syntactic considerations operating among the NA constructions to be found in the language. Adjectives (if they are that) such as 'elect' in 'president(-)elect' are instances of the lexical criteria, and often represent set expressions or a somewhat formal register, eg. 'the body proper'. Adjectives such as 'present', 'concerned' and 'affected' have different meanings before and after the noun:

present members	(ie. who are members <u>at</u> present)
members present	(eg. at a meeting)
concerned people	(cf. thinking people)
people concerned	(cf. people affected, see below)
affected people	
people affected	

The NA construction may often be paraphrased, eg. 'such people as are (at present) affected' and this semantic distinction comes out clearly in the example given by Quirk et al (1972) of 'visible stars' versus 'stars visible': the latter is clearly restrictive, and as they observe represents a temporary situation, though the basic meaning of the preposed form is retained. It is not clear whether only certain lexical items permit this construction - especially -<u>able/ible</u> adjectives - or whether any adjective of a particular type of meaning permits it. The syntactic criteria appear in examples of complex AdjPs in indefinite NPs:

> Soldiers timid and cowardly don't fight well.
> (Quirk et al (1972)).

An important group of adjectives, which I have not classified above are adjectives such as <u>poor</u>, <u>wretched</u> and <u>dear</u>, indicating attitude of speaker. They occur before the noun in French (<u>pauvre</u>, <u>méchant</u> and <u>cher</u>) but show a difference in meaning post-nominally. Such expressions, denoting attitude of speaker, are also to be found pre-nominally in Cemuhi, Welsh and JACALTEC, all prepositional VS languages, normally NA. In Finnish however,[4] which is postpositional SV and normally AN, certain such words occur after the noun, including:

> parka, poor kulta, dear raiska, wretched

This phenomenon of changing the order to denote attitude of speaker is very apparent when looking at minority adjective ordering, though it is not clear from the above examples if it correlates with Pr/Po or with VS/SV. In addition we must bear in mind its overlap with other semantic phenomena: their number is closely identified with groups D, E and F, eg. old, small, nice, when expressing affection, scorn or admiration. Furthermore their use is generally appositive rather than restrictive. In my present study of adjectives, however, my task with respect to them will be firstly to note their ordering when it arises, but more importantly, to isolate their effect from that of other semantic considerations.

6.2 Evidence from the survey on ordering

0.

In Greenberg's work on word- and morpheme-order (1963), his 19th Universal was as follows:

> When the general rule is that the descriptive adjective follows, there may be a minority of adjectives which usually precede, but when the general rule is that descriptive adjectives precede, there are no exceptions.

In this section I shall first of all consider NA languages with a minority of preceding adjectives, and then consider the question of whether any symmetries may be found in the AN languages - thus contradicting Universal 19 - and also in the NA~AN languages available.

In the present sample, the following NA languages have among their descriptive adjectives a minority AN order:

> Welsh, Cemuhi, Tiv, Masai, JACALTEC, ZAPOTEC, SULKA (Pr), Abkhaz and WASKIA (Po).

There are of course plenty of NA languages with preceding indefinite or numeral adjectives not listed above. These will be commented on separately in the following sub-sections.

Among the AN languages Finnish and English have already been examined for a minority NA order: Finnish is in the sample. In addition we find AN languages such as Chamorro placing material (K) and nationality (L) items after the noun (see below). These languages suggest at least a partial exception to Greenberg's hypothesis.

In the following four sub-sections I shall survey the evidence for what type of adjectives are likely to be found among the minority groups. I shall first take NA languages, but on account of the clear differences between the Pr and Po languages in the type of adjectives preceding, I shall treat these two groups separately.

6.2.1 Pr/NA languages

In this section I shall investigate tendencies in Pr/NA languages to deviate from their dominant NA order.

For the purpose of this section I shall list the 23 languages found in this sample to be Pr/NA:

Cemuhi, Fijian

Berber, Maori, Mâsai, Niue, Welsh, Yapese,
JACALTEC, ZAPOTEC

Cairene, Fulani, Kinyarwanda, Malay, Margi,
Thai, Tiv, Vietnamese, Yoruba, Zulu, OLO, SULKA

Iraqw

All but five prepose something before the noun.

The languages underlined are those I shall pick out for comment. Cemuhi, Mâsai, SULKA, ZAPOTEC, Tiv and Welsh have a minority of descriptive adjectives preceding the noun. It may be noted that a number of these and other Pr/NA languages also have preceding indefinite (A), quantitative (B) or ordinal (C) expressions. Of these I shall pick out Niue, Maori and Fijian for discussion.

The remaining two underlined languages, Yapese and Kinyarwanda have one AN item each which I shall discuss, but their status as adjectives is questionable.

In considering languages with a minority of AN adjectives, I shall first examine Cemuhi, Mâsai, SULKA and ZAPOTEC, ie. those with the fewest preceding adjectives, and after summarising the patterns of these, I shall look at JACALTEC, Tiv and Welsh, where a stronger argument could be put forward for a semantic rather than lexical criterion in preposing adjectives.

For Cemuhi, Rivierre (1980) identifies four groups of items which may stand between article and noun:

(i)	cuŏ	small, minor
	pwŏkŏ	large, major
	ukèi(u)	old, former (both, it appears, pejorative)
	culi	little (affection)
(ii)	cèi(u)	one (and other numerals)
(iii)	pètaapwŏ	separately
(iv)	ju	real (often prefixed to give a set phrase meaning)

Rivierre appears not to identify adjectives as such in Cemuhi: the words for 'big' and 'small' are termed 'appreciatifs', while NA

'adjectives' are in effect either stative verbs or locative expressions:

bo	magat	'red hat'	(Cemuhi)
hat	is=red		
a	mwà	wéaga	'the Weaga clan'
Det	clan	W.	

Note that Cemuhi is a VOS language, so that normally we may distinguish the attributive and verbal uses of stative verbs by their position after the noun or before the noun phrase. The third and fourth words listed under (i) are in fact statives which may precede when the connotation is pejorative or affectionate.

Masai permits kutí, a few, ọtí, small, pookí(n), all, to precede or follow[5]:

	áyíéú	ọl-	ọtí	alẹm	'I want a small sword'
	I=want	Msg	small	sword	
or	áyíéú	ọl-	álẹ́m	ọtí	

Note the gender prefix to the first element of the NA or AN complex. This is also found with likaí, another, which must precede. Another indefinite, họọ, any, also precedes, but replaces gender-marking.

In SULKA the general pattern is one where the determiner is repeated before noun and adjective (Müller (1916)):

a	kọri	a	vankér	'heavy stone'
Det	stone	Det	heavy	

or much less frequently:

a vankér a kọri.

With few exceptions the adjectives are derived from intransitive verbs. However five adjectives,

ke, small (E)
mi, usual (A?D?)
vu, many (B)
sim, good (F)
ker, bad (F)

may only precede the noun, and the article is not repeated:

a	sim	blom	'a good woman'
Det	good	woman	

a mi yok 'an ordinary (piece of) taro'
Det usual taro

The last four may also be used as adverbs.

In ZAPOTEC adjectives or 'statives' may be regarded as stative verbs, which are normally added as the second part of a compound with nouns[6]:

gi- šNyâ 'red flower'
flower red

komíd- yähL 'cool food'
food cool

cf. nayähL komîd, the food is cool, byähL komîd, the food was cooled.

The statives rob, big, and wên, good, are listed in Briggs (1961) but not exemplified, and there are additionally items which may only occur in compound postnominally, such as -rŏ, big, and -gaí, male. It is not clear whether these should be regarded as a separate syntactic class, or as statives which like some adjectives in English cannot occur predicatively.

In addition to the statives discussed above, there are five, according to Briggs, which may precede the noun in compound: primeer, first, LuLt, last, maal, bad, ganaz, purely, nothing else but, and za, clear, settled (of coffee). Thus we have ganáž-byûž, just children. In addition there is the prefix s-, with tone perturbation on the noun, with the meaning 'other' (morpheme 541):

s- té- šmân 'another week'
other one week

Now the criteria for the placing of these adjectives is clearly lexical: a limited number of lexical items qualify, even though there is some semantic pattern[7]:

A.	Indefinite	all (Mâsai)
		another (Mâsai)
		separately (Cemuhi)
B.	Quantitative	many (SULKA)
		a few (Mâsai)
		Numerals (Cemuhi, ZAPOTEC, SULKA)
C.	Ordinal	first (ZAPOTEC)
		last (ZAPOTEC)

D.	Modal/Tense/Age	usual, ordinary (SULKA) (D?)[8] real (Cemuhi) nothing but (ZAPOTEC) old, former (Cemuhi - Value?)
E.	Dimension	big (Cemuhi) small (SULKA, Cemuhi, Mâsai)
F.	Value	good (SULKA) bad (SULKA, ZAPOTEC) dear little (Cemuhi)
G-H.	Speed, Colour	(none)
I.	Physical Property	settled, clear (ZAPOTEC)
J,K,L.		(none)

Underlined are items occurring in more than one language.

It may well be that in each of these languages we have a handful of idiosyncratic constructions, as with 'galore' in 'drinks galore' in English. Note that in ZAPOTEC, where only those specifically mentioned by Briggs have been listed, three - primeer, LuLt (cf. el último) and maal - look suspiciously like Spanish.

Notwithstanding, the AN adjectives listed occur, with one or perhaps two exceptions, in the first seven categories A-G. As mentioned above, the languages so far exemplified appear to have a limited number of lexical items preceding the noun, possibly in idiosyncratic constructions. We may now look at three more languages, JACALTEC, Tiv and Welsh, each with a larger number of AN adjectives, and each suggesting more clearly criteria by which we may expect adjectives to precede or follow.

I shall first turn to JACALTEC.[8a] According to Craig (1977) adjectives in general follow the noun. We thus have examples such as

xawil	te'	Hah	ac	tu' ?
you=see	Cl	house	new	that

'Do you see that new house?'

Such items had already been considered by Day (1973), but here they are analysed as adjectival predicates in a relative clause:

how	no'	cheh	'The horse is mean/vicious'
is=mean	Cl	horse	

no' cheh≠how 'The horse which is mean/vicious'

where the juncture sign ≠ indicates lengthening of the preceding vowel /consonant and joins a relative clause to its antecedent. By contrast the adjectives which precede are clearly identifiable either semantically or morphologically, suggesting that JACALTEC should be treated as NA, with the relativised adjectival predicate being the most general construction.

The distinction between NA and AN constructions is set out in Craig's work, and she identifies in particular a group expressing 'admiration, scorn or fondness'. Such adjectives are

	ich	old
	niman	big
	tz'ulic	cute
	xil	used, worn out
and	leje	confused.

Thus the word niman, big, may occur either before the noun, as in

ch'en	ich	niman	ch'en	tu'
Cl	old	big	rock	that

'that big old rock'

or after, as in

s-ti'	ha'	niman	lit. 'the mouth of the
its mouth	water	big	big water'

translated by Craig rather more loosely as 'the river bank'. Presumably the latter example does not imply either admiration, scorn or affection. We may note the parallel between JACALTEC and Cemuhi, where two of the AN words, one meaning old or former and the other meaning little, are preposed when used pejoratively or to signify affection.

As well as the above semantic group, Craig identifies a large morphological group taking the suffix -la in prenominal position. She identifies two major sub-groups, of which the first denote colour (H). Such adjectives can in fact occur post-nominally, presumably as a relativised predicative adjective, but when they occur in pre-nominal position, they may take on a derived meaning:

saj-la	camiẍe	'white shirt'
white	shirt	

but saj-la ha' 'clear water'
water

We may also note that colour adjectives referring to animals ending in -iH may either precede or follow.

Another group which may precede with -la are the 'positional adjectives' those which end in the derivative suffix -an, and often associated with physical condition or shape (I):

hopan- la camiɨe 'torn shirt'
torn shirt

Other adjectives are given by Craig, and we have three examples: tz'il(-la), dirty, chacan(-la), deaf, and how(-la), which last I shall now focus upon: this adjective has already been exemplified above in the NA form no' cheh how, the mean or vicious horse, but in the AN example

how-la tx'i'
dog

we have the more specific meaning of rabid dog.

In Tiv, we find that most adjectival concepts are represented by NR relative clauses. There are however a limited number of adjectives of which the following precede (Abraham (1940)):

(i) Diminutive/augmentative

wan, dzəgə̱,
small big

plus others in this group,

(ii) other

kɔkɔ, gɔgɔ, bo, gbougbou,
long long ugly big/fat

It is a characteristic of these adjectives that they change the class of the noun following them. Thus we have the word wəgh, of the kun class:

wəgh kun 'this hand'
hand this

where classes of nouns are named by the form of the demonstrative following them. The characteristic of this class is the suffixation

of -gh. After an adjective, however, nouns go into Class I, or Class XI for the plural, regardless of their original class:

wəgh,	wan	wə
hand	small	hand
iyoɣgh,	kɔkɔ	you
yam	long	yam

Note the dropping of prefix and/or suffix on the above nouns, when they go into (Abraham's) Class I.

It appears that bo may follow with the meaning 'bad'. Other adjectives which may precede or follow are

tamən,	ir,	tsə
great	black	old

When these follow they go into the genitive and the original class of the noun is retained:

bua	u	tamən	'big cow'
cow	of	big	

Adjectives which must follow are ndor, moist, cold, gən, any, other, kɔ, raw, unripe, hə, new, tsutswa, first:

m-	kə	-m	ma	kɔ	-m	'unripe pepper'
X	pepper	X	of	raw	X	

where X is the characteristic affix of the man- or Collective Class (X).

These adjectives too go into the genitive after the noun to which they are attributive: they are according to Abraham, remains of derived verbal nouns (which he terms 'True Verbal Nouns'), so that they really mean 'pepper of unripeness', etc. The genitival link may be dropped for gən, any, other, for tsə, old and in certain set expressions, hə, new.

Note additionally that such genitival adjectives still bear the class suffix of the noun they modify, a form of partial class agreement.

We must also contrast these adjectival genitives with nominal genitives such as

dzwa	uu	Tiv	'the Tiv language'
language	of	T.	

where 'Tiv' shows no gender agreement.

Thus we see that Tiv is a nominal genitive and translates what would be a nationality adjective in English (group L).

We may therefore list as AN examples in Tiv:

D	*	tsə	old (but not hə, new)
E	*	tamən	big
		wan	small
		kɔkɔ	long
		gbougbou	big, fat
F		bo	ugly (but not the words for 'good', 'bad')
J	*	ir	black

where the asterisked words may follow.

What we thus find in Tiv is two very clear groups, the Augmentative/Diminutives, plus the relative clause construction. In between we find the device of placing the verbal noun in the genitive, from which an adjectival NA construction has emerged, though even here we find that certain of these NA words for 'ugly', 'great', 'black' and 'old' may also precede the noun.

We may now turn to Welsh and here we find that though grammar books will list a limited number of AN adjectives, including, say, hen, old, prif, chief, unig, sole (as opposed to 'alone' when it follows), we can in fact list them easily under classes A (indefinite), B (quantitative), C (ordinal), D(i) (modal/tense) plus hen (under D(ii)):

A.	rhyw	some (sg.)		
	pob	every		
	y holl	all the		
B.	llawer	many	rhai	few
	rhyw	some (pl.)		
	plus	numerals		
C.	prif	chief		
	unig	sole		
	hoff	favourite		
	plus	ordinals, comparatives and superlatives		
D(i)	gwir	real	gau	false
	tragwyddol	complete	rhannol	partial
	union	exact		

(continued)

	plus possibly	under D(i)		
	rhesymol	reasonable	cam	wrong
	ddifrifol	serious		
D(ii)	hen	old		

In addition we have the adjectives annwyl, dear, parchedig, revered, which seem to go in the same group of attitudinal adjectives that we encountered in French, Finnish and possibly Cemuhi and JACALTEC.

Whilst we may easily list the above adjectives under A, B, C, D(i), and the attitudinal adjectives, it is not at all clear whether this is simply a motivation for lexical specifications, or whether we have an on-going set of semantic criteria.

The analysis so far has concentrated upon expressions which were classifiable as adjectives. Thus Yapese and Kinyarwanda have been missed - Yapese (NA) preposes the item chi, small, which seems to be a syntactic class on its own, and Kinyarwanda (also NA) preposes -ndi, other, which may well be a determiner, since it cannot co-occur before the noun with a demonstrative. (ZAPOTEC and Mâsai have already been mentioned in connection with preposed items for '(an)other'.)

The languages so far considered are PrNA languages with a limited number of AN constructions, partly determined semantically, but mainly determined lexically, though there appears to be a semantic motivation: groups A to F show a more or less similar propensity to prepose their most general adjectives, except that big and small from E appear to be the strongest if we discount the Welsh adjectives under D(i) (tense/modal), as may be seen in Table 6.1. (See next page)

We thus have evidence from one group of six PrNA adjectives to the effect that E should come higher than F on the hierarchy, and that F should come higher than groups G to L. For D, we find that old preposes more often than good or bad in F (Value), but whether D should come higher than E is uncertain: there are a large number of preposing adjectives under D, but they are all concentrated on Welsh, where a semantic rather than lexical criterion may well be at work.

It might be thought that D and E prepose more often than groups A, B and C, since we may observe that of these six languages Cemuhi postposes 'all', while only Welsh, Tiv and Yapese prepose it. (Mâsai has both orders and the facts for ZAPOTEC are unclear.) Similarly we find

TABLE 6.1

PREPOSING OF SOME COMMON ADJECTIVAL CONCEPTS

(PrNA languages)

old	small	big	bad	good
Welsh	SULKA	Tiv	SULKA	SULKA
Tiv	Tiv	Cemuhi	ZAPOTEC	
Cemuhi	Cemuhi			
	Mâsai			
	Yapese			
(3)	(5) (2)		(2) (1)	
D(ii)	E		F	
Age	Dimension		Value	

that Mâsai and Tiv postpose numerals and three of them, Cemuhi, Welsh and Yapese, prepose them. These figures might suggest that 'all' and the numerals are roughly equivalent on the hierarchy to D and perhaps E. If however we look at other Pr/NA languages which have not yet been selected for preposing descriptive adjectives, we find a considerable number of such languages have been observed to prepose A and B items.

I shall therefore look more closely now at groups A and B, and consider the treatment of cardinal numerals and the translation of 'all'. Of the 23 PrNA languages six postpose 'all'. These are Cemuhi, Fijian, Maori, Niue, Thai and SULKA. Another four have both postposing and preposing constructions: Berber, Kinyarwanda, Zulu and Mâsai, whilst information was inconclusive for ZAPOTEC and OLO.

Slightly more languages postposed numerals, so that PrNA languages were evenly divided between NumN and NNum. Here I have only considered indefinite numeral expressions, ie. excluding those translatable by 'the three men'. Mâsai (VSO), Margi, Zulu, Kinyarwanda, Tiv, Fulani, Yoruba, Thai, OLO (SVO) and the only SOVPr language Iraqw had NNum, ten in all, while Malay and SULKA had both NumN and NNum.

I shall now exemplify the above pattern, singling out three languages for special discussion: Niue, Maori and Fijian. These VS languages of

the Austronesian grouping appear to break the hierarchy by postposing 'all' (A) but not numerals (B). In fact in these three languages the appropriate word may be 'floated' to a post-verbal position, yielding a construction similar to the Quantifier Floating transformation posited for English. In particular this has been discussed for Niue in Seiter (1980), where he shows that oti, all, may occur either post-verbally or post-nominally (see 'Quantifier Float'):

Moua e maua mo Sione e tau mata afi oti
get Erg we with S. Abs Pl match all

'Sione and I have won all the matches'

or: Moua oti e maua mo Sione e tau mata afi.
(where case-marking varies before pronouns and propernames, and before common-noun NPs)

Examples may also be found in Fijian[9] and Maori of 'floating' from a post-nominal position. 'Unfloated' examples are

keimami ā rogoca na veikā kece ka baleti keimami
we Pa hear Det thing all concerning us

na gone na Viti (Fijian)
Fijians

'we have heard all the things concerning us Fijians'

ka tuku- na ngā whare katoa ki te ahi (Maori)
Pa put Pass the(pl) house all to the fire

'all the houses were burnt'

'Floated' examples are

era moce kece tiko 'they are all asleep' (Fijian)
they're asleep all now

ka tuku- (na) katoa- tia ngā whare ki te ahi (Maori)
Pa put Pass all Pass the houses to the fire

'all the houses were burnt'

Note that katoa in the second example has floated to a position between verb and passive suffix, optionally repeated on the verb. This example was given to me by Bruce Biggs (Auckland), to whom I am indebted for a valuable discussion of Maori syntax.

The ability of quantifiers and even numerals to 'float' from or out of their NPs will be further illustrated in subsequent sections.

Normally when a preposition precedes the noun, the numeral in a NumN language will stand between preposition and noun. This does not happen however if the quantifier floats, precedes the preposition or in some cases obliterates a preposition. This last occurs in Niue, where the absolutive case preposition may drop before a numeral:

mate	tuai	(e)	ua e	kuli	'two dogs have died'
die	Perf	(Abs)	two	dog	

(Such obliteration also occurs in certain expressions in English: thus an Indian restaurant notice stating 'This restaurant is open on every Sunday' is ungrammatical, even though 'on Sunday' is all right.)

We may therefore conclude for PrNA languages, if we consider only 'all' and numerals from A and B, that 'all' should come slightly higher on the hierarchy than cardinal numerals, which in turn come higher than D and E: nine languages were found with NumN order, preposing numerals, a larger number than those preposing 'small' and 'big' (five and three respectively).

Two further points may be noted about the data observed: the first is that ordering criteria among items A to D(i) tended to be syntactic and sometimes semantic (eg. Welsh adjectives under D(i)), while criteria among items D(ii) to F tended to be lexical, resulting perhaps in a much sharper tailing off down the hierarchy. The second point to be made is that a correlation is observable between NumN and VS or verb-initial languages. If we were to take the hierarchy alone into account, we would get the prediction that such items in Pr languages tend either to stand between preposition and noun or at least stand on the same side of the noun as the preposition. The correlation with VS however suggests that such items also have a tendency to stand between verb and subject noun or nominal expression, something which can clearly be associated with the patterns of quantifier and numeral floating that I have already illustrated above.

6.2.2 Po/NA

In this section I shall investigate tendencies in Po/NA languages to deviate from their dominant NA order. The Po/NA languages may be listed as follows:

Ewe, Lugbara, GUARANI

Abkhaz, Bawm, Burmese, Djingili, Kanuri, Mandingo, Nubian, Wati, GUAYMI, HIDATSA, POMO(E).

DAGA, KATE, NASIOI, SENTANI, SIROI, WASKIA.

There are 20 languages listed: underlined are those I have singled out for comment. Except for missing data, the other languages postpose everything discussed in this section.

We may now turn to the postpositional languages which have a minority of preceding adjectives: Abkhaz and WASKIA. Examples to be found of nationality (L) are:

wəy	à-	kərtwa	jɣab	(Abkhaz)
that	Det	Georgian	girl	
	'that Georgian girl'			

Takia	kadi	kuareng-kuareng	pamu	(WASKIA)
T.	man	old (pl.)	this	
	'those old Takia men'			

In the relevant grammatical references, these preceding expressions have been described as preceding adjectives.

Abkhaz also permits expressions of material to precede (K):

à-	kʼərmətʼ(-tʷʼə)	yʷnə		(Abkhaz)
Det	brick	of?	house	
	'a/the brick house'			

Although this construction has not been described in Hewitt (1979) as a preceding adjective, it nevertheless looks identical to the nationality construction, and we may perhaps say that its status as an adjective is as debatable as the status in English of 'brick' as an adjective.

We also find preceding expressions of nationality in Burmese and GUAYMI:

bəma	sapei	'Burmese literature'	(Burmese)
Burma	literature		

ni	Oriko-	bu	slota-	ye	(GUAYMI)
a(n)	Tikobian	?of	child	Dat	
	'a Tikobian child'				

where the -bu is always (and only?) found with such nationality

expressions in GUAYMI.

Burmese is an interesting language to consider: we may identify at least three methods of rendering adjectival concepts. The group I have identified as adjectival for the purpose of comparison is the verbal noun, following the described noun in close juncture. As an example we may take the verb kàuñ, be good. The associated verbal noun has the prefix ă-, to give ăkaùñ. We then add this to the 'head' noun in close juncture:

ămyoù-	kaùñ	'good family'
family	good	

(Two morphophonemic processes may be observed: the dropping of the prefix ă- before a non-initial element of a compound, and the voicing, indicated by underlining, of certain consonants after certain tones in the previous element of the compound.)

We may exemplify this construction more fully as follows:

D(ii)	(age)	myoú-	haùñ	'old city'
E	(dimension)	ămyi?-	hyei	'long root'
F	(value)	ămyoù-	kaùñ	'good family'
H	(colour)	ăhma?-	mè	'black mark'
I	(physical)	yei-	pu	'hot water'

We can in fact extend the above NA pattern upwards in the hierarchy to groups A and B, but the post-modifiers have wider syntactic properties than the above 'adjectives', and they are referred to in Okell (1969) as 'noun auxiliaries'[10]:

A:	ăhkàñ-	taìñ-	hma	'in every room'
	room	every	in	
B:	ngăyou?kaùñ-	hkălèi		'a little pepper'
	pepper	little		

However other constructions do occur, in which adjectival concepts precede. Certain of the verbal nouns may either precede or follow when they take loose juncture:

eìñçi	ăpya	or	ăpya eìñçi
shirt	blue		

where the two morphophonemic processes outlined above have not occurred. Furthermore these 'adjectives' may precede the noun in the original

verbal form, as part of a relative clause:

kaùñ-	té	lu	'good person'
be=good		person	

where the -té suffix is added to relative rather than main clauses. Compare the above with

hʂoù-	té	lu	'evil/bad person'

The above pattern suggests that a somewhat different set of adjectives will precede the noun in pre- and postpositional languages: in prepositional languages they range from A to F with perhaps colours as well (H). In postpositional languages they are nationality (L) or material (K). Note that in French, which is Pr, adjectives of colour (H), material (K) and nationality (L), are most likely to follow.

I shall now turn to groups A and B.

If we look at Indefinite items (A), we find that clear data is not available on the 'all' construction for five languages (Djingili, GUARANI, GUAYMI, HIDATSA and POMO (E)). Otherwise all but Kanuri and Ewe place the word for 'all' after the noun.

Kanuri preposes the 'all expression, while Ewe repeats the noun both before and after siáa, all, as in the following example[11]:

lã	siáa	lã	'every animal'
animal	every		

In GUARANI the words for 'every' are entéro and opá. The former, a Spanish borrowing, precedes the noun[12]:

entéro	ma'é	i-	varàto-	vé	raka'é
every	thing	3obj	cheap	-er	formerly

'everything was formerly cheaper'

No examples were found for opá. A floating construction did however occur, with the suffix -pá/má suffixed to the verb:

o-	ma'ẽ-	má	umĩ	hente	ore-rehé	
3subj	look	all	those	people	1pl	at

'all those people looked at us'

(with -má occurring after nasal phonemes).

It was observed in section 6.1 that if a Po language was NA, it was

overwhelmingly likely to postpose numerals and the qualifier 'all', in line with their placing higher up on the hierarchy than descriptive adjectives. The statistics for numeral position - with all 20 PoNA languages except GUARANI (NumN), Abkhaz (NumN~NNum) and Djingili (no data) showing a NNum pattern - illustrate the importance of an unbiassed sample: in Greenberg's sample there were seven PoNA languages, of which two appeared to have NumN order (Basque and Burmese) and Guaraní was listed as having both orders. Seven languages in a geographically biassed sample are simply insufficient for establishing either a hierarchy or its actual presence, let alone when twelve rungs are contemplated.

It is generally the case, as far as data is available, that the numeric expression stands between noun and postposition, but this does not happen in Mandingo[13]:

kɛ -	kɔ rɔ̀ -	niin-	nu	tɛ ma	fíla
man	old	little	Pl	between	two

'between two little old men'

The hyphens between the words for 'man', 'old' and 'little' indicate Mandingo's form of close juncture, which determine the tone contours within the phrase: the last syllable of the penultimate element of the close compounding bears the accent.

The data on PoNA languages elucidated above appears to be in accordance with the hierarchy in that only two languages appear to violate it: 'all' is higher up the hierarchy than descriptive adjectives, and it should therefore always follow in PoNA languages, but we find that in Kanuri 'all' precedes. Similarly numerals (B) should follow, but in GUARANI they precede. Strictly speaking then, we do not yet have evidence that these items are at the top of the hierarchy for postpositional (Po) languages, or at least above descriptive adjectives. We could refute this in the next section if we found that numerals and 'all' precede in all PoAN languages. They would then be in a roughly similar position on the hierarchy. In fact we find that this is not the case.

The evidence on K (Materials) and L (Nationality) has in fact been clearer in placing K and L at the bottom of the hierarchy: a number of PoNA languages pre-posed expressions of material and/or nationality, and as we shall see, no corresponding number of PoAN languages cancel

out the tendency by placing such expressions after the noun.

To this extent we appear to have the same hierarchy for Pr and Po languages: the pattern among the Pr languages is that items at the top of the hierarchy, especially A to E or F, tend to precede, with K and L items following. The reverse occurs with the Po languages. What we cannot do, however, is use postpositional languages - so far only NA languages but as will be seen in the next section also AN languages - to identify variations in ordering for the descriptive adjective in the middle of the hierarchy.

6.2.3. AN

For the purpose of this section I shall list again the AN languages in the present sample, this time classified by Pr/Po:

Pr: CAYUVAVA, SHUSWAP

Chamorro, Greek, Hausa, Mandarin, Maung, Tiwi, YUROK

Po: Bardi

Finnish, ZOQUE

Burushaski, Hindi, Ijo, Japanese, Kannada, Nama, Newari, Turkish, CARIB, MAIDU QUECHUA, FORE, MARIND

(There are 25 languages listed, 9 are Pr and 16 Po.)

The languages underlined have been selected for comment.

As far as I am aware, only Finnish postposes any specific group of adjectives, which I have exemplified in section §6.1 above. We must however note that these are very marginal to Finnish, and since they suspend case-marking on the noun Whitney (1956) argues that they are not really adjectives, but elements in a noun-noun compound:

Ei	poika	para-	lla	ole	rahaa
does=not	boy	poor	All	be	money(Prt)

'The poor boy has no money'

However it is not possible to deduce much from the suspension of case-agreement, despite the discussion of it in chapter 4, on account of the exceptional exclamatory nature of the construction.

The above is the nearest we have in AN languages to items with some

syntactic or lexical status as adjectives, but which violated the normal AN ordering constraint. If however we wish to extend the patterns illustrated for Pr/NA and Po/NA, we may further consider whether expressions with the meaning of K (Material) and L (Nationality) have any tendency to follow in Pr/AN languages, and whether any similar expressions with the meaning of A to F or G have any tendency to follow in Po/AN languages.

I shall first look at Pr/AN languages, and I shall look closely at Hausa, which has only a limited number of so-called adjectives, using other constructions, such as 'of X-ness' to express the semantic concepts under consideration.

The situation in Hausa is the most straightforward, and the number of strictly adjectival items is limited, although it may in fact be argued that they are really nouns governing a genitive. Hall and Kirk-Greene (1973) list the following which I list here according to my own classification:

D(ii)	sābō	new	tsōhō	old
E	bàbba	big	ƙàramī	small
			ƙànƙanè	small
	dōgō		gàjērē	short
F	-		mūgù	bad
H	farī	white	baƙī	black
	jā	red		
	kōrè	(light) green		
	shūdī	(light) blue		
	rāwayà	yellow		

Dixon (1977) lists a similar number, but includes ƒanyē, fresh, unripe (I), may be added.

Use of these adjectives may be exemplified by the following[14]:

bàbba-	n	gidā	'(the) large home'
large	of	home	
tsōhō-	n	mùtûm	'(the) old man'
old	of	man	

Note that tsōhō may also be used for 'former', D(i):

tsōhō-	n	firāministà	'(the) former PM'
old	of	PM	

We may also identify preceding indefinite expressions (A or B?):

wani	irì-	n	tsuntsū		'a certain kind of bird'
certain	kind	of	bird		
kōwàcè	kàsuwā				'every market'
every	market				
waʃansu	mânya-	n	jirāge-	n	samà
some(pl)	big(pl)	of	boats	of	sky

'some big aeroplanes'

Here the genitival construction is not used.

Examples have been given above from A, B?, D, E, F (bad only) and H, of which the last four make use of a genitival construction. However such genitival constructions may be replaced by NA, though such examples occur less frequently, and their use is essentially, it appears, for indefinite constructions:

gidā	bàbba	'(a) big house'
house	big	
yārò	ʃànʃanè	'(a) small boy'
boy	small	

It must be noted, however, that when only a dozen or so 'adjectives' are listed for Hausa, this is not so much a restricted syntactic class, as a restricted lexical class. Thus while we may express 'round' (I) and 'angry' (J) by means of the mài construction:

I	tēbùr	mài	kèwayā	'round table'
	table	with	go=around	
J	mùtûm	mài	fushī	'angry man'
	man	with	anger	

derived adjectives are also available:

kèwàyayye-	n	tēbùr
round	of	table
mafùsàci-	n	mùtûm
angry	of	man

We may therefore observe that

(a) groups A, B?, D, E, F (bad only), H, I and J make use of an AN ordering, in which the noun often comes second as a genitive,

(b) groups B, F, G, H, I make use of the mài construction:

B	kudī	māsu	yawā	'much money'
	money	with(f)	abundance	
F	aikì	mài	kyâu	'good work'
	work	with	goodness	
G	kùnkuru	mài	sauri	'fast tortoise'
	tortoise	with	speed	

(c) groups C (ordinal), K, L also make use of the genitival construction, only here it is the noun that acts as head:

C	yāro	na	biyu	'the second boy'
	boy	of	two	
K	kwàle-kwàle-	n	kātākō	'wooden boat'
	boat	of	wood	
L	ɗalìbī-	n	bàtūrè	'English student'
	student	of	Englishman	(of English or from England)

In addition dukà, all (A), and the cardinal numerals (B) also follow their noun, though without using a genitival construction.

Another AN language where certain adjectival concepts follow is Chamorro. First of all consider the following examples, where the 'adjective' comes first[15]:

B	meggai na bi'shi	many times
E	i dánkolo na taotao	the big man
	i dikike' na patgon	the small child
H	i betde na kareta	the green cat (cf. Spanish verde)

The item straight before the connector na is in fact a stative verb which when used in an attributive construction must precede the noun.

On the other hand we have the following K and L examples:

K	i papet	aseru	sandpaper
	(paper)	(steel)	
L	i gime'	Filipino	Filipino house
	i relós	Amerikanu	American watch

In such examples, however, the last items are nominal post-modifiers and thus have an entirely different origin to the stative verb pre-modifiers.

I shall now turn to postpositional AN languages, where we might expect, in mirror image to Pr/NA languages, that certain adjectival concepts in English are translated by post-nominal items. There is indeed such evidence, but it is to be found exclusively in groups A to E, with barely anything that may be described as descriptive adjectives.

I shall first of all consider MARIND, where certain adjectival concepts may follow[16].

Under B we find that quantitative expressions may precede or follow with a difference in meaning:

aha	inah	'two houses'
house	two	
inah	aha	'the two houses' (or, of course, 'both houses')

Also the word otiv, many, may precede or follow, so that apparently both anim otiv and otiv anim mean 'many people'. However they also appear to mean 'everybody' ('the many people'?, A), especially when followed by ipe, the.

MARIND, according to Drabbe (1955), is limited on numeric ordinals (C). However comparative and superlative expressions may follow the noun: this is achieved by adding the element ha or isi to the adjective, which may then precede or follow its noun:

sok-vakra	arera^{i}	ha
knife	sharp	real

'a really sharp knife', 'a sharper knife'
or 'the sharpest knife'

according to context.

The meaning of ha, real, takes us to D, in that it may follow the noun on its own:

anem hà, a real man, really a man.

Turning to Japanese we also find following A and B expressions[17]:

tamago o top kowasimasita
egg Acc ten broke

'(I) broke ten eggs'

hoteru no heya o mina soozi simasita
hotel 's room Acc all clean finish

'(I) cleaned out all the hotel rooms'

Note that the bracketed 'I' may be replaced by any subject pronoun. Note also that the second example is ambiguous: it could also mean 'everybody cleaned the hotel rooms'. The postposing of indefinite and quantitative expressions certainly has exceptions:

san- bon no enpitu o ka- tte imasu
three Cl 's pencil Acc buy- ing be

'(He) is buying (the) three pencils.'

including an ambi-fixal construction:

dono ike ni mo san- syurui no sakana ga ita
which pond in also three kinds 's fish Nom were

'In every pond were three (different) kinds of fish.'

As with the Mandingo example given earlier, postposed indefinite and numeric items follow any case-marking, though in the case of mo, the markers wa, ga and o (topic, nominative and accusative) get obliterated.

The choice between NumN and NNum in Japanese depends partly upon which item is to be given greater emphasis, the numeric or nominal expression: the second item bears the greater emphasis. But two further points should be observed: firstly, if we have an expression such as 'the three pencils' representing a more or less closely linked group, the numeric expression must precede, and secondly, if we have a topicalised expression, the position of the numeral depends upon whether or not it is part of the topic.

Emphasis is also a criterion in FORE as well, in that the postposing of the numeral gives it a greater prominence and is thus considered in Scott (1978) to be the more marked position:

na:má 'tara- bi mpintáwé 'they are in two houses'
house two in they=are

Note that in FORE, unlike Japanese, the postposition must follow the numeric expression.

In Ijo the position of the numeral depends upon the number itself: multiples of ten tend to follow while others precede[18]:

déin-	a	suei	'thirty nights'
night	Link	30	
ói,	ma	burú	fìnì 'twelve yams'
ten	two	yam	plus

Here only the 'plus' gets postposed. However we find with Ijo that 'all' follows the noun, as occurs also in Burushaski:

béi	warì-	mọ-	sẹ	'all these houses'	(Ijo)
this	house	the	all		
ho·l	uyo·n-	ɛ	haha	ɛtʊmʌn	(Burushaski)
army	all	Erg	charge	did	

'the whole army charged'

Amongst the languages exemplified above, primarily for indefinite (A) and quantitative (B) expressions, little tendency was observed to postpose descriptive expressions. This was only observed in MARIND, where the item ha, real, was found. The only other Po/AN language observed to postpone such items was CARIB, where two items were noted, with the meanings 'late' (D?) and 'small' (E)[19]:

ya:wo	mi	'(my) late uncle'
uncle	late	
mo:ro	waya:rímbo	'mɛ 'the little basket'
the	basket	small

We must note however that these two items behave in a similar syntactic manner to possessed nouns, in that they cannot be separated from the noun they follow, and represent the only thing that can stand between a possessed noun and its possessor noun. However I do not propose to assign any dominant ordering to the placing of 'small', since CARIB does in fact have two 'ordinary' adjectives for 'small', kowa:ro and imembo(ko:).

If we now turn to A and B modifiers of the noun, in particular 'all' and the numerals, with respect to both PrAN and PoAN languages, we should expect, according to the hierarchy, that Pr languages will over-

whelmingly pre-pose A and B items, while no prediction is made for PoAN languages: in fact among the 9 PrAN languages all but Hausa prepose numerals (B), while all but YUROK prepose 'all' (A). Data is unavailable on 'all' for Maung and Tiwi[20].

Among PoAN languages we might expect a larger number of PoAN languages to postpose A and B items than PrAN: there are in fact 16 PoAN languages in the sample, of which four, Ijo, Japanese, FORE and MARIND, may postpose numerals (no data on MAIDU and Bardi). Among the 12 for which data is available for 'all' (data missing or unclear for Bardi, MAIDU[21], ZOQUE and FORE), there were two, Ijo and Burushaski, which postpose 'all', whilst another three, Nama, Japanese, and MARIND, place an expression for 'all' before or after the noun. Thus among the Po languages there is a greater propensity to postpose A and B items than among Pr languages: furthermore among Po languages there is a slightly greater tendency to postpose A items than B items, and this is in accordance with the hierarchy, at least for 'all', though perhaps a larger sample would be desirable to check this.

We may note in conclusion for AN languages that the evidence for the hierarchy was much less clear than for NA languages. Among NA languages, the position of all and of numerals strongly correlated with Pr/Po, whereas among AN languages both tended to precede - though among Po languages there were a number of languages in which they optionally followed, sometimes following the postposition as well.

It was not easy among AN languages to find examples of descriptive adjectives being postposed, at least in the sample. This would appear to be in accordance with Greenberg's Universal 19. Nevertheless when such postposing did occur, it did appear to display a mirror image pattern to that of NA languages: the hypothesis of the hierarchy would require the postposing of D, E and perhaps F among postpositional languages. This was to be found only to a limited extent: in MARIND, the word for real (D) could be postposed, and in CARIB the words for late (D?) and small (E) also followed the noun. There was also a clear pattern, in accordance with the hierarchy, in which expressions of material (K) and nationality (L) were more likely to follow in Pr languages than other adjectival expressions. This did not occur with Po/AN languages.

6.2.4. NA~AN

The last group I shall look at is one of seven languages for which a dominant order was difficult to establish. These were

Palauan, Sango, Tagalog, Tsou (Pr)
Ngandi, LUISENO, PURO (Po)

The pattern observed above for prepositional languages was confirmed. In particular 'all' and numerals were preposed, with the one exception of Sango. In addition Tsou showed some variation in its expression of its descriptive adjectival concepts: normally the adjectival expression may either precede or follow its noun, but in either case they are separated by ci. However the words okó, small, and méoi, big, take another particle, no, which is also the one used before a genitival expression. Another group precede the noun without any separating particle:

	A.	acɯha	all
		maezo	same
	D.	oá	not
		zóu	real
	D?	nia	late, ancient
other		iachí	one's own
		toehúŋu	common

Sango is also a language which will have to be looked at in greater detail. Samarin (1967) lists a set of ante-nominal and post-nominal 'adjunctives', of which I shall now consider those classifiable under A-G:

	Ante-nominal		Post-nominal	
A	gí	only	ngá	also
			vɛ́ní	Intensive (cf. meme)
			ndé	different
			tongasó	like that
			kóé	all, completely
B	mbéní	some	míngí	many
				(plus numerals)
C	kózo	first		

(continued)	Ante-nominal		Post-nominal	
D	táa/ʔgí	real		
	finí	new		
	ngbéré	old		
	mbakɔrɔ	old		
	(vieux	old)		
E	= kótá	big	= kótá	big
	= kété	small	= kété	small
	ndurú	short, near		
	yɔngɔ́rɔ	distant		
F	nzoní	good		
	sioní	bad		
	= pendere	beautiful	= pendere	beautiful
	= séngé	mere	= séngé	mere
G	-			

The four adjectives preceded by = may precede or follow the noun and they are therefore listed in both columns, opposite each other. Note the French borrowings: apparently French numerals, both cardinal (B) and ordinal (C) may precede the noun, as may vieux/vieille and même, the latter apparently with the meaning 'even'.

In the above list, groups A and B seem to go partly against the hypothesised hierarchy. Among the remaining groups, we find the following ante-nominals among H, I and J:

H	vɔkɔ́	black	(also NA)
	vurú	white	
	bingbá	brown	
I	kpíngba	solid	
	ngangó	hard	
J	búbá	stupid	

The above list, of ante- and postnominals covering groups A to J, with the exception of G (speed), leaves uncovered a number of concepts expressed by adjectives in English. These are mainly expressed by genitival constructions.

yíngɔ́	tí	basánze	'home-made salt' (L?)
salt	of	wild	
kpí	tí	sindi	'sesame paste' (K)
yángá	tí	Sango	'the Sango language' (L)
mouth	of	S.	
mbéní	yí tí	saleté	'something dirty' (K?)
some	thing of	dirty	

No examples are given of expressions such as 'fast car', 'slow tortoise' (G), though adverbial expressions of speed are certainly to be found.

For the three postpositional languages data is somewhat limited: no data is available on 'all' for PIRO and Ngandi, while in LUISENO the adjective coun is used for 'all', and presumably may therefore either precede or follow its noun.

It is not clear whether numerals precede or follow in Ngandi, but in the other two languages, they appear to precede, although the hierarchy would suggest that they should either follow, or follow adjectives in coming either before or after the noun.

In conclusion we may say for these seven languages with alternative NA and AN orderings that the hierarchy appears to be strongly approximated to among the Pr languages. Thus we found preposed words denoting real (D(i)) in Tsou and Sango, as well as in the Pr/NA languages Cemuhi and Welsh. Similarly small precedes in Tsou, as well as in the Pr/NA languages Cemuhi, Tiv, Yapese, SULKA and (optionally) Mâsai. The word denoting old precedes in Sango and Tsou - where it also means late - as well as the Pr/NA language Tiv, and indeed in Cemuhi, JACALTEC and Welsh, especially to denote scorn or affection. These examples are particularly worthy of note, in that although there are only four such languages in the sample - Pr with both NA and AN ordering, they bring out the hierarchy more clearly than the 23 Pr/NA languages. These observations contrast with the Po languages with both NA and AN ordering, which like postpositional languages generally, showed much less evidence for the adjectival hierarchy.

6.3 Conclusion

I shall now attempt to summarise the results of the survey carried out in this present chapter. It must be borne in mind in the findings set out below that the evidence for the hierarchy put forward is much clearer among Pr languages, than among Po languages, and additionally much clearer among NA languages than among AN. This consideration applies especially amongst the descriptive adjectives examined.

Taking the top of the hierarchy first, we found that among Pr languages, the propensity for 'all' and for numerals to precede was greater than the propensity for descriptive adjectives to precede (in this sample indeed the latter tend to follow). We may see this by tabulating allN versus Nall, the position of 'all' before or after its noun:

Table 6.2

(Pr)	allN	either	Nall	Total available	Not available
NA	9	4	8	21	2
AN	7	-	2	9	2

We see along the first row that among NA languages allN and Nall are almost evenly balanced, while along the second row AN languages are overwhelmingly allN.

A similar picture occurs for numeral position, as may be seen in the following table:

Table 6.3

(Pr)	NumN	either	NNum	Total available
NA	10	3	10	23
AN	9	-	2	11

where it is only among NA languages that any number of languages postpose numerals.

Turning to postpositional languages we find a similar pattern in reverse. If we take 'all', we find that it is only among AN languages that any number of languages prepose 'all':

(for Table 6.4 see next page)

Table 6.4

(Po)	allN	either	Nall	Total available	Not available
NA	1	1	12	14	5
AN	7	3	2	12	4

Similarly for numeral positions, it is only among AN languages that any number prepose numerals:

Table 6.5

(Po)	NumN	either	NNum	Total available	Not available
NA	1	1	16	18	1
AN	10	4	-	14	2

Turning now to descriptive adjectives, I shall first set aside the Material (K) and Nationality (L) groups. Among Pr/NA languages, an attempt was made to distinguish between semantically and lexically determined preposed adjectives. Under the semantic criteria, the D(i) group was clearly identifiable in Welsh, with words denoting real, false, and former, for example, while a group of Diminutive/Augmentatives was identifiable in Cemuhi and Tiv. Otherwise small, bad and old were most frequently identified as preposing under lexical criteria. Among Po/AN languages, evidence was much more limited for postposing items, and here there was only evidence for lexical criteria, with items noted in MARIND and CARIB, denoting small, real and late (deceased).

Coming to the bottom of the hierarchy, clear evidence was found of a tendency for K and L items to follow in prepositional languages and precede in postpositional languages.

It is most important to note that among descriptive adjectives, minority items in terms of ordering are rarely found in more than 10% of the languages concerned. The only exception to this was words denoting small, which preposed in five out of the 23 NA languages. But what we have in fact shown in this chapter is that Greenberg's Universal 19, mentioning minority-order adjectives in NA languages, is roughly correct as a statistical generalisation. However the minority adjectives which precede in prepositional languages are clearly of a different type to those which precede in postpositional languages.

There is of course evidence for a mirror-image pattern amongst AN languages, but it is much more limited.

This is an important conclusion, for which an explanation will be sought in the next chapter, and an important device in bringing the evidence out was the hypothesised adjectival hierarchy. This has of course been the prime role of the hierarchy, and any more ambitious employment of it would require a much larger sample, able for example to distinguish more clearly between statistics of 5% and 10%. In addition we would have to be rather more careful to identify how far Pr/Po was the critical conditioning factor: thus numeral position may well display a greater correlation with VS/SV than with Pr/Po, and it is entirely plausible that the position of material expressions (K) has a greater correlation with NG/GN than with Pr/Po.

7. Adjuncts.

0.

The discussion of the last chapter attempted to pave the way for a semantic account of adjectives by explaining the ordering patterns of various types of nominal 'modifiers' distinguished semantically.

The first task of this section will be to show that the evidence of the above data is to refute the Operator-Operand account Vennemann and Bartsch put forward in their Natural Serialisation Principle. That will be the task of the first section of this chapter where it will be seen that their departures from Montague's approach, particularly in treating object NPs as operators, do not succeed in solving the problems they set out to solve.

At one stage of their discussion however, in Vennemann (1973), use is made of an appositional construction to solve certain problems in the ordering of determiners and numerals, and I shall discuss in sections 7.2 to 7.4 the relative merits of exploiting this device - since we have it - and of finding ways of treating adjectives as either operators or operands depending upon the order used in the languages concerned.

Finally in section 7.5 I shall argue, primarily from English, that in situations where an appositional construction seems attractive, a separate concept of an adjunct seems more desirable: perhaps the relationship of an adjunct to apposition is similar to that of a subordinate clause to a co-ordinate clause. In any case it enables us to make use of expressions such as 'complete victor', where 'complete' describes not the noun to which it is adjoined, but one implicit in it, ie. 'victory'.

7.1 Problems with Natural Serialisation

The above discussion has attempted to pave the way for a semantic account of adjectives by explaining the ordering patterns of various types of nominal 'modifiers' distinguished semantically. The task of this section will be to show that the evidence of the data of chapter 6 refutes the Operator-Operand account Vennemann and Bartsch put forward in their Natural Serialisation Principle.

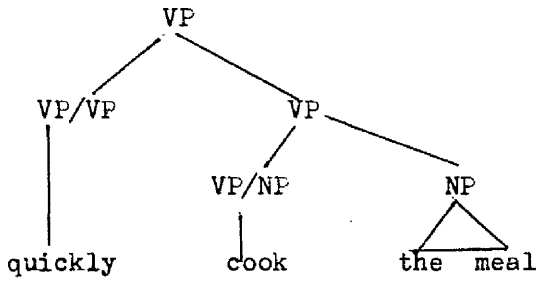
Natural serialisation is a state in which 'operators' in a language consistently precede or consistently follow their 'operands' and examples of such operators and operands, divided into two groups A and B for later discussion, are as follows:

	Operators	Operands
A.	Object NP	Transitive verb
	Complement NP	Pre/postposition
	Genitive NP	Noun (possessed)
	Non-finite verb	Modal/auxiliary
B.	Adverbial phrase	Verb
	Noun 'modifiers' (incl. determiner, numeral, adjective, etc.) excluding genitives	Noun

This approach, like that of Montague, differs from the earlier phrase structure approaches, and resembles the structures of logical form, in distinguishing something like function and argument. However Vennemann and Bartsch use the terms 'operator' and 'operand' to cover not only functions such as 'square of' and 'father of' but also predicate-term relationships and 'modal operator-proposition' relationships (see Vennemann (1974)).

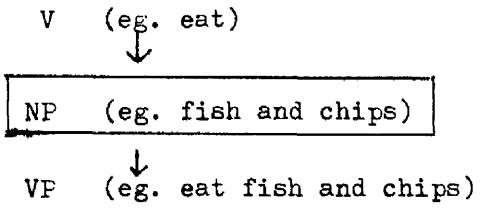
We may now consider the difference between Montague's approach and that of Vennemann. There is no difference for items listed under B, in the sense that both approaches treat them as functions or operators. The difference lies in the other pairs under A, and this is best illustrated with the combination of transitive verb and object NP.

In the Montague approach, the verb (VP/NP) operates upon the object NP to give a verb phrase, in contrast to an adverbial phrase which operates upon a VP to give a larger VP (VP/VP):



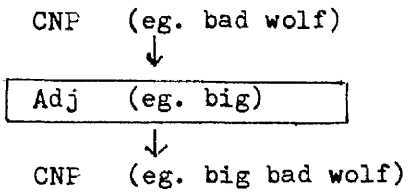
By contrast all verbal modifiers in the Vennemann/Bartsch approach are operators, except perhaps the subject. As a result the prediction is made that among VO languages, where operators follow their operands, adverbial modifiers of the verb will also follow the verb, and that among OV languages, where operators precede their operands, adverbial modifiers will also precede. Such a prediction is in fact partly borne out by Greenberg's Universal 7, where 'rigid' SOV languages (those which cannot put the object after the verb) may not postpose adverbial phrases.

We may now turn to what Vennemann and Bartsch term 'Natural Generative Grammar' which they present as a development of Montague's programme. It was felt that there was uncertainty as to what was operator and what was operand in a syntactic construction: a defining criterion was found in the form of <u>category constancy</u>. We thus now find that object NPs are treated as operators, in contrast to Montague's original 'mistake' of treating the transitive verb as operator:



What has happened here is that the NP operates upon a transitive verb to produce a verb phrase. Another example is that of an adjective which operates upon a CNP to produce a larger CNP:

(see next page)



The above criteria might suggest that word-order universal patterns may be determined purely from syntactic criteria, particularly in the case of Vennemann and Bartsch's work, since their criteria appear to produce better predictions in the area of adjectival and adverbial modifiers, and in any case Montague was not concerned with word-order as such. However both approaches took semantics into account, as I shall now recount.

Looking first at the work of Montague, we have in effect a constraint on the interaction of semantics and syntax, in that the operators of his programme lend themselves to intensional logic: they may take as operands or arguments things which either do not exist or are incorrectly referred to, yet lead to true non-negated sentences. We may see this in the following examples, where the intensional operator or function is underlined:

They are hunting unicorns
They are talking about unicorns.
That is a picture of a unicorn.
I saw the supposed unicorn.

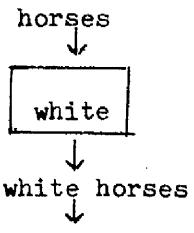
Here we have examples of intensional 'operators' appearing in VO, Pr, NG and AN constructions, all with operator before the imaginary or incorrectly described 'operand'. Other intensional uses of adjectives are of course those categorised above under E and F: a good thief, a small elephant. We appear here to have an inconsistent pattern in that AN should be consistent with Pr, VO and NG, when in fact this is not the case: however the evidence of section 6.2 has been that among adjectival concepts, the ones that most lend themselves to analysis as intensional operators are most likely to be AN among languages with Pr and NG ordering.

The work of Bartsch and Vennemann is also very much concerned with the semantics of adverbs and adjectives and their interaction with syntax. Bartsch's work on Adverbialsemantik (1972, tr. 1976) examines German adverbial expressions in considerable detail, both their relative

ordering and translation into logical form. Her work on the Semantics and Syntax of Number and Numerals (1973) does in fact break into an area where Montague has published little. Both Bartsch and Vennemann consider in detail 'relative adjectives' (eg. 'small' in 'small elephant', ie. of a size smaller than the norm for an elephant) in their work on Semantic Structures (1972) and Vennemann's work on Explanation in Syntax (1973) develops this in his consideration of demonstrative, numeral and descriptive adjectives. Now both writers accept Montague's treatment of adverbs and adjectives as intensional operators and basically go further into the formal definition of such operators.

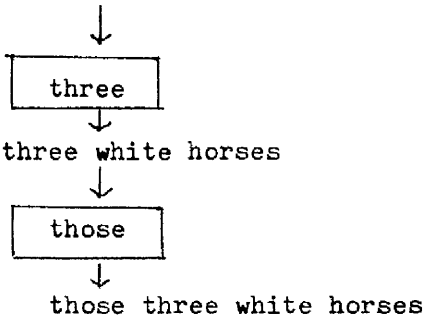
Now there are two points we may note with reference to their semantic analysis. Firstly we may note for adjectives such as 'red' or 'round' for which an extensional approach is possible, that the associated operators are defined so that expressions such as 'red car' involve the intersection of the set of cars and the set of red objects, and thus, at the semantic level, the relation between 'red' and 'car' is completely symmetric. There are then no a priori grounds for treating one rather than the other as operator: it is only to maintain uniformity of treatment within the syntactic class of adjectives that we treat such adjectives as operators rather than as operand. The difficulty occurs however in that when we treat adjectives as operators, we find that the non-intersectional adjectives are less likely rather than more likely to observe the Natural Serialisation principle, behaving in effect as if they were less likely to be treated as operators. Such a result is totally incongruous with the Bartsch/ Vennemann approach, let alone that of Montague.

We may secondly observe an incongruity in the treatment of determiners, which is at first sight specific to only a limited number of languages: in Vennemann (1972) demonstrative, numeral and descriptive adjectives are all treated as operators. Normally the descriptive adjectives operate upon the noun, then the numerals operate on the resulting expression, and finally the demonstratives are brought in, giving us expressions such as:



(continued)

(continued)



This means that adjectives must come closer to the noun than numerals do, and numerals must come closer to the noun than demonstratives. Problems now occur with languages referred to in Greenberg's 20th Universal, such as Kikuyu, which according to Greenberg (1963) use the order noun-demonstrative-numeral-adjective[1]:

ng'ombe	iris	ciakwa	ithatũ	njega
cows	those	of-me	three	good

'those three good cows of mine'

(marking for the n-class underlined)

Once we get more than one noun modifier with such an order the process of evaluation in terms of operator and operand breaks down, since operators such as the demonstrative stand between parts of its operand expression. To deal with this Vennemann makes use of an appositional relationship:

[[N] [Dem [Num [A]]]]

Now in fact this won't do, in that the last two items, the numeral and adjective, may indeed also occur in the order adjective-numeral: furthermore we find in Swahili, adjacent both genetically and geographically, that although its unemphatic order has noun-adjective-numeral-demonstrative, it can nevertheless vary the order for purposes of emphasis. But it is important to note that Vennemann is forced to introduce an additional modifier concept which is neither operator nor operand, represented in his notation as

[[.....] [.....]].

Having noted the problem of non-intersectional adjectives - and orderings such as that of Kikuyu - for the Natural Serialisation Principle, we may turn to the problem of what can be done about it. The Bartsch/Vennemann scheme may be summarised as follows:

(i) Adjectives, including demonstrative and numeral, are all operators, though appositional constructions are required on occasion.

(ii) Genitival expressions, direct objects and prepositional objects are likewise treated as operators, permitting a pattern of word-order consistency.

(iii) Circularity in the above scheme is avoided by means of the principle of category constancy.

Now a straightforward way of solving the problem of non-intersectional adjectives is to take item (ii) above and treat the direct object NPs, etc. as operands rather than as operators. This would be very welcome, at least in that we may now revert to the approach of Montague - and indeed my own approach of the first chapter. There are however two difficulties. The first is that the analysis of intersectional adjectives must be examined more closely: not only are there no a priori grounds for treating them as either operators or operands, but also the statistical data, in particular the stratal analysis of chapter 2, suggests that a separate analysis is required.

The second point we must note is that much of the above arguments about possible semantic structures depends heavily upon the evidence of word-order statistics: while this is far from being a bad thing in itself, we must bear in mind that we have had to abandon the Bartsch/Vennemann principle of category constancy. Ultimately, if a return to circularity is to be avoided, we must find something to put in its place.

The task of this chapter however will be to look more closely at the adjective-noun relationship. Is it sufficient to say that certain adjectives may be indifferently treated as operators or operands, or must we bring in something like apposition? In my next three sections I shall examine in turn the analysis of adjectives as operators, appositional items and then as operands. In the final section however I shall introduce a fourth structure, which I shall term adjunction, which I believe provides the answer to the main problems considered in this work for adjectives, and indeed many adverbial expressions.

7.2 Adjectives as operators

It might be thought that the situation in which adjectives could be most easily treated as operators is that where the ordering for a language is Pr/NG/AN or Po/GN/NA. We then get a very straightforward scheme:

Operator	Operand
Pre/postposition	NP
Head noun	Genitive NP
Adjective	Head noun

In considering this question I shall first consider the example of English, which has Pr/AN plus NG as the most frequent genitival construction. We then have the hypothesis that an adjective-noun construction has the form:

$$[\text{red} \quad [\text{car}]]$$

This approach would appear to be much the most straightforward, since there is nothing to stop us from treating all adjectives in English as operators, as do both Montague and Vennemann. However questions do arise. In particular we may note that the approach suggests that the attributive form of the adjective is basic, and the predicative construction somehow derived from it, since a predicative adjective has no noun to operate upon. Advocates of this type of approach will propose that such an adjective will have a dummy argument, so that 'is small' comes out as

$$[\text{is} \quad [\text{small} \ [\Delta]]]$$

Essentially the expression comes out as 'is a small one'. There are, of course difficulties with this approach, and two may be immediately cited: firstly we have to get rid of the determiner when the head noun is removed or replaced by a dummy, nor indeed can we take expressions such as 'is a nice small one' and derive 'is nice small'. But secondly and more importantly, even if we do accept an analysis such as

$$[\text{is} \quad [\text{small} \ [\Delta]]]$$

there is nothing to stop us representing 'small elephant' as

$$[\,[\text{small} \ [\Delta]][\text{elephant}]\,]$$

where the attributive adjective has an appositional relationship with its noun. If we pursue this line, where [[X] [Y]] has the meaning 'X which is Y', or vice versa, we are very close to the 'WHIZ-deletion' of Transformational Grammar, where the similar syntactic properties of attributive and predicative adjectives are captured: indeed an appositional structure at the semantic level is not incompatible with WHIZ-deletion at a syntactic level. It is not at all clear however whether there is any point in proposing such a structure for non-intersective adjectives which are never predicative anyway.

In fact a very straightforward approach suggests itself when we realise that not all adjectives in English may be predicative. We may see this from the following examples, partly taken from Quirk et al (1972):

This argument is { the main one. / *main }

This is the only opportunity.

This opportunity is { the only one. / *only. }

Note that in some cases, failure to use the 'a/the....one' form will elicit a different meaning of the adjective:

My friend is an old one.

My friend is old.

We thus have a number of adjectives, where a predicative expression can only be formed using 'a/the....one', if even that may be used:

*My wife is a former one.

*This end is the very one.

*That fool is an utter one.

and again different meanings may be elicited:

He is a big baby (ie. very babyish)

That baby is a big one.

This discussion leads us to a hypothesis that in English only those adjectives that cannot be used predicatively should be treated as operators:

[very [end]]

[bottom [candidate]]

[former [president]]

They might include others such as 'particular', among the 69 adjectives listed by Siegel in her (1980) work as having both intensional and extensional readings, which may be predicative for only certain meanings:

Our charlady is particular.
? Our problems are particular.

It is important to note that I have used the expression 'only those adjectives...' rather than 'all those adjectives that cannot be used predicatively'. Thus the adjective 'wildcat' cannot be used predicatively - if indeed it is an adjective - but can we really say that 'bottom', 'former' and 'wildcat' come in the same class of intensional adjectives in the way that Siegel suggests? Perhaps the first two are operators by virtue of being ordinal (C) and temporal/modal (D(i)), whilst 'wildcat' is a noun taking part in a noun-noun construction.

It is important to note that the 'only' implies that we do not include adjectives such as 'false' and 'partial', which may in fact be used predicatively:

That beard is false,
Their victory was partial,

even though we learn from the above sentences that the beard referred to is not a beard, nor the victory a victory, so that in Montague terms, these adjectives are intensional.

If we are to admit of something like an appositional relationship for those adjectives that may be used predicatively, presumably we can formulate structures something like:

[[false] [beard]],
[[partial] [victory]],
[[white] [rabbit]].

(Here I have dropped the Δ for convenience, since there will be no longer any situation in which its presence and absence stand in contrast.)

I shall now attempt to list groups of adjectives which would be eligible for treatment as operators:

(i) Adjectives with ordinal content, such as
chief, principal, ultimate
which cannot be predicative:

* This success of his was ultimate.

(cf. this success of his was his ultimate one).

I am not clear about

top, middle, bottom

since some predicative expressions are more acceptable than others:

* People who read the Times are top.

? Martin was top (of the class).

Martin came top.

If the second example is questionable in adult language, it is certainly used by younger school-children. In any case the use of top in the 'Times' example is different from the 'classroom' examples.

We would thus have

[top [people]].

I am not clear how 'main' should be classified: its behaviour is very similar to that of other adjectives above, but it seems to be more superlative than ordinal, though in the previous chapter I have in fact treated superlatives together with ordinals under C.

(ii) Adjectives such as former and future:

* That PM is former.

* That PM is future.

These come perhaps under tense (D), next in line to (C), though I shall presently show that other items such as the adverb sadly appear to behave similarly. Certain adjectives which would easily have been classified under D(i) do not come into this group:

That beard is false. (as mentioned above)

That roof is temporary.

A property of some at least of these adjectives is that they have corresponding adverbial forms which go with adjectival expressions:

a formerly excellent student

is still a student, in contrast with

a former excellent student

who is not.

Such adverbials cannot take adverbial phrases:

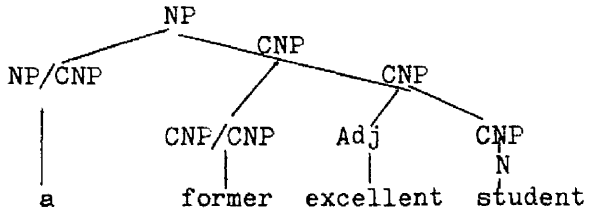
a formerly sadly disappointing player

as opposed to

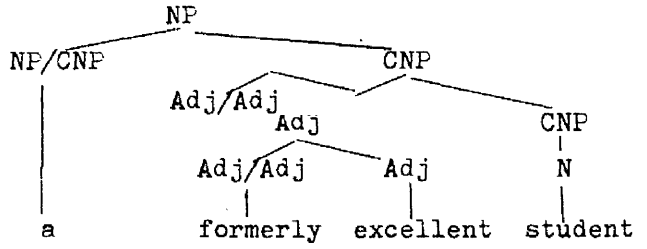
a very sadly disappointing player.

In the first example of the two, the student used to be sadly disappointing as a player: it does not mean that he is still disappointing, but we are no longer sad about it. In the second example, 'very' goes only with 'sadly'.

We may express this structurally by



as opposed to



where the node marked as Adj and its relationship within the higher CNP has still to be examined. (It would in fact be possible at this stage, though unrevealing, to treat it as CNP//CNP.)

We cannot extend this group, so far only exemplified by temporal adjectives, to adjectives such as <u>apparent</u> for two reasons: firstly, though we cannot say:

* That criminal was apparent,

we can however say

That man's crime was apparent.

Secondly we might think that this is only acceptable when not only is the person concerned apparently a criminal: he actually is a criminal. But we may also have

His crime was only apparent.

Similarly we find an example using alleged where the adjective may occur predicatively when qualified:

His guilt was alleged by the prosecution,

without implying that the accused was actually guilty. The reverse situation occurs with attributive uses, as with

His very apparent guilt took us somewhat aback,

where on account of the 'very' the guilt of the man concerned is implied. But on no account can either apparent or alleged be predicated to the person having such properties, as opposed to the properties themselves.

However we can extend this group to adjectives (and corresponding adverbs) which might be called sentence-modifying. Here we have sad and sadly:

His sad failure disappointed us all.
His sadly disappointing failure puzzled us.

Note that here also

His sadly very disappointing failure vexed us,

shows that sadly takes adjectival rather than adverbial expressions. Thus we find that

* His sadly apparently disappointing failure...,

won't do, as opposed to

His sadly apparent failure/disappointment...

It is possible that in this group we should only include the adverbs, rather than both adverbs and adjectives: but note that though we can say

His frankly rude behaviour....

to indicate the frankness of the epithet-user,

His frank rudeness...

can only refer to the frankness of the rude person.

Above is a tentative outline of the adjectives as operators in English. It may indeed be the case that all attributive adjectives could be treated thus in English: however I have concentrated upon the example of a limited number of types where we do not need to provide a

mechanism to allow for predicative adjectives. I shall now consider the example of adjectives in other Pr/AN languages - and of Po/NA. I shall focus here primarily upon adjectives which appear to partake of a genitival construction. In the present sample the only clear examples of this type appear to be Tsou and Hausa, which we may exemplify as follows[2]:

gida-	n	sarkī	'the chief's house' (Hausa)
house	of	chief	
bàbba-	n	gidā	'the large house'
large (one)	of?	house	
to	féou	no	úa	'the skin of the/a deer' (Tsou)
Det	skin	Det	deer	
meoi	no	cóu	'a big man'
big	Det	man	

(where the Det's vary according to spatial remoteness and in the above forms precede non-topicalised NPs)

It would seem a straightforward matter in the two Hausa examples to treat them both as operator-operand structures:

[gidan [sarkī]] (house of the chief)

[bàbban [gidā]] (large house)

where the first word in each example acts as operator. The difficulty occurs when there is a separate linking element, such as the no in Tsou, the of in English, and indeed the uncontracted form of the linking element in Hausa:

gidā	na	sarkī
house	of	chief

Note that gidā and na sarkī may each function separately as NPs, the former meaning 'house' and the latter meaning 'that of the chief' or 'the chief's (one)'. It would then seem reasonable to treat the two expressions as coming together in apposition:

[[gidā] [na sarkī]].

Note that an adjectival noun such as bàbba may also be placed after the noun, and here there is no difficulty about treating such expressions

as nouns in apposition:

[[gidā] [bàbba]]
house large

No link is required when the adjectival noun follows rather than precedes.

Thus to indicate the same concept we might use both an operator-operand construction and an appositional one. But it must be noted that the adjectival noun as an operator differs morphologically from when it functions appositionally, in contracting with the linking na. We then have something which is semantically equivalent to apposition when the genitive denotes identity, the big thing which is a house, but differs when the construction denotes possession, ie. the big thing belonging to the house.

This ambiguity of meaning can be localised in English to the variety of meanings of the item of, and presumably we can do the same with the linking element in Hausa. The ability to analyse adjectival constructions in this way becomes less apparent in the other type of construction I cited, where both adjectival and genitival constructions employ simple juxtaposition. This situation arises in a number of Po/GN/NA languages. In such languages we would expect operators to follow operands, to give a structure

[[Operand] Operator]

An example we may cite is that of GUARANI, where the genitive occurs straight before the head noun[3]:

aróye kóta 'the bank of a stream'
stream bank

These juxtaposed nouns cannot be in apposition, since 'stream' and 'bank' refer to different things. Here it would be appropriate to have an operator-operand relationship:

[[aróye] kóta]

with the operand first. In the case of an adjectival construction, the 'adjective' follows:

petei ivitú mareté 'a strong wind'
one wind strong

However we do not have a syntactic similarity as in Hausa, since these 'adjectives' are in fact stative verbs bare of inflections. Consequently they correspond to the potentially predicative adjectives of English and could for similar reasons be treated as having an appositional relationship:

[[ɨvɨtú] [mareté]]
wind strong

though we cannot exclude the possibility without a greater familiarity with GUARANI, that they cou[ ] as operators:

[[[ɨvɨtú] mareté]]

in line with both the Montague and Vennemann/Bartsch approach.

It might be thought that there will be no great difficulty in a Pr/AN or Po/NA language of treating adjectives as operators: after all the adjectives concerned include intensional adjectives. However such an approach provides no valuable insights. It is more revealing in English for example to dispense altogether with the concomitant mechanism whereby predicative adjectives have a dummy operand, since the remaining adjectives which cannot occur predicatively are adjectives such as main, chief and former, which lend themselves very readily to an analysis as operators.

It is also revealing in English not to treat other adjectives as operators, since we then capture the fact that certain adjectives, such as partial, incomplete and false, which are strong candidates for an intensional analysis, are nevertheless perfectly acceptable in predicative expressions such as

That beard is false,

where any dummy operand, if necessary, occurs identically in predicative and attributive expressions!

The above reservations do not preclude us from treating adjectives as operators in languages such as Hausa and Tsou, where adjectives resemble head nouns in genitive constructions. Nevertheless a case has emerged for treating adjectives as something like appositional items, and this is what I shall consider in the next section.

7.3 The appositional relationship

I shall now focus more directly upon the question of whether we can treat adjectives, or perhaps adjectival nouns, as appositional items. An important alternative is of course one in which adjectives are operators or operands, whichever maintains word-order consistency: consequently alternatives must be considered to apposition. To do this I shall consider three types of construction where apposition might be considered, and I shall try to show in these constructions, that they become increasingly more difficult to deal with without some concept such as apposition.

The first type we should consider is the juxtaposition of nouns. We have already exemplified Hausa in this regard, in the last section. Further examples may be found in Burmese and PIRO[4]:

ăpya ein̆ci 'blue shirt' (Burmese)
or ein̆ci ăpya

sato kwanonur-potu sutxo (PIRO)
one poor very woman
'a very poor woman'

In the constructions of both languages the equivalent of our adjectives is a noun, so that we have two nouns juxtaposed. In the Burmese example the same adjectival concept occurs either before or after the noun, though it is not clear that items other than nouns of colour can do this. For PIRO it appears that a modifying noun may precede or follow its head noun.

It is this ability to place adjectival concepts before or after the associated substantival concepts that distinguishes these examples from those exemplified in the last section, taken from Pr/AN and Po/NA languages. The problem there was whether adjectives were operators or items in an appositional relationship. The question with the languages exemplified so far in this section is whether something like an appositional relationship holds or whether instead the semantics is such that it doesn't matter which is operator or operand. In this latter case we may choose whichever is consistent with the overal operator-operand ordering of the language.

This problem of identifying the relationship may be further illustrated when we look at expressions which are not simply juxtaposed, as in

Tagalog, where a linking item intervenes:

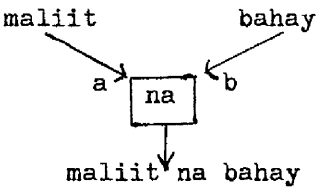
maliit	na	bahay	}	'small house'
small	+	house		
bahay na maliit				

- with the na becoming -ng after a vowel.

There is more than one way in which we can analyse the above expressions. One is to treat na/-ng as an infixed binary operator:

maliit na bahay

which has two inputs and an output of the same two items in apposition:



However this is not the only possible analysis for

X na Y

One could for example employ an analysis of the form

X na Y

where the na is a dummy operator. Consequently the evidence from Tagalog is at present no clearer than that of say PIRO where simple juxtaposition occurs. In fact I shall suggest in section 7.5 that the na/-ng indicates a semantic region:

bahay na maliit 'small house'

or where there is a particle such as ang, ng or sa preceding, marking topicality or grammatical function:

ang bahay na maliit. '(as for the) small house'

(where ng is not to be confused with -ng, the alternant of na).

I shall now turn to the third construction where an appositional relationship may occur. This is one where either case agreement occurs, or the preposition or postposition is repeated for a number of items in the noun or pre/postpositional phrase. The best example of this appears to be Maori, where there is no clear distinction between prepositions and case-markers.

If two NPs are in apposition in Maori, any preposition governing them is repeated:

ma	tōna	tupuna,	ma	Pau
for	his	grandfather	for	P.

'for his grandfather Pau'

The repetition does not occur when nouns are juxtaposed in a non-appositional construction:

o	era	whare	raupo
of	those	house	(bulrush)

'of those raupo (bulrush) houses'

Such juxtapositions of nouns rather than NPs occurs when the second noun denotes material or purpose.

It could be argued that when prepositions are repeated as in the above examples, we do not have nouns or NPs in apposition, but prepositional phrases in apposition: if something is for my grandfather Pau, it is for my grandfather and for Pau. Further evidence for such an analysis occurs when we note that conjunction of NPs following a preposition is effected by repeating that preposition[5]:

i	whakatokia	te	mara	ki	te	uwhi,	ki	te	taro,
Past	be-planted	the	field	with	the	uwhi	with	the	taro

ki	te	kumara
with	the	(sweet potato)

'the field was planted with uwhi, taro and kumara'

We may not be able to place the 'uwhi', 'taro' and 'kumara' in apposition, but we can place adverbial or prepositional phrases containing them in apposition, since they jointly describe the manner of the planting. We may therefore ask whether languages showing case-agreement between noun and adjective have an operator-operand relationship between adjective and noun, or whether perhaps we have instead two phrases in apposition, perhaps even adverbial phrases in apposition.

It will be noted again that apposition in these examples has cropped up in a new guise: can these prepositional phrases be regarded as being in apposition or partaking in an operator-operand relationship where the most 'convenient' item is chosen as operator?

In this section, and indeed the previous one, I have attempted to show

that a concept of apposition may usefully be set up, and applied in a consistent fashion to two or more phrases which jointly describe something to which they refer directly. Thus in the Maori expression

rua	kumara	'kumara pit'
pit	(sweet-potato)	

we might be entitled to say that the words 'rua' and 'kumara' jointly describe a pit for (cooking) kumara, but only 'rua' refers to it directly. This is not therefore an example of apposition: the phrase cited above for 'his grandfather Pau' on the other hand is, since 'his grandfather' and 'Pau' both refer directly to the person concerned. The only difficulty that now arises is that even though we have such a consistent concept, do we wish to treat it as such, and if so, do we wish to subsume it under the operator-operand concept, or treat it as something separate?

Before we can discuss this directly however, it is necessary to consider one last issue: that of adjectives as operands.

7.4 Adjectives as operands

I shall now look at adjectives as operands, and in so doing I shall pay particular attention to Pr/NA and Po/AN languages. The main point I shall consider is that of why certain types of adjective, which have the strongest tendency to precede in prepositional languages and to follow in postpositional languages, suggesting that they should be treated as operators, should nevertheless lend themselves in certain languages to treatment as operands.

I have already discussed the idea that if an adjective can partake of an appositional relationship it can also be treated as either operator or operand. However I would like to consider situations where a treatment of the adjectives concerned as operands might be preferable.

The example that immediately springs to mind is that of Maori, where I have already illustrated the non-appositional juxtaposing of nouns:

whare	kowhata	'stone house'
house	stone	
rua	kumara	'pit for (cooking) kumara'
pit	(sweet potato)	

Williams (1923) states that nouns placed after other nouns in this way become adjectives. However we also find that when adjectives stand on their own, eg. pai, good, they do not refer to the object so described, as in Hausa, but to the abstract quality:

te	pai	'goodness'
the	good(ness)	

We would then have an alternative analysis, in which an expression such as

he	whare	pai	'a good house'
a	house	good(ness)	

would in fact mean 'a house of goodness'.

A further piece of evidence for this supposition is that only one adjective may stand after the noun in this way, necessitating an appositional expression:

he	kowhata	nui,	he	mea	taimaha
a	stone	large (ness)	a	thing	heavy (ness)

'a large heavy stone'

It would be nonsense under my analysis to have:

* he kowhata nui taimaha

since this would juxtapose abstract nouns either to put them in apposition or with the meaning 'a stone of heavy largeness'.

The question now arises as to what we do with relative clauses. In Maori a clause is added to the antecedent noun (with an 'adjective' optionally inserted), a deictic particle is placed after the verb, replacing any post-verbal imperfective particle:

te	kakahu	e	whatua	na	e	koe
the	garment	Ipf	be-woven	there	by	you

'the garment which you are weaving'

(note the passive, or 'Inverted Construction' see Williams (1923), § 75).

Presumably the corresponding main clause would be:

e	whatua	ana	e	koe
Ipf	be-woven	Ipf	by	you

'(it) is being woven by you, you are weaving it'

If we place the relative clause in apposition with its antecedent it means that the clause directly refers to the antecedent. If on the other hand we treat it as an operand we could consider the clause to refer to the fact: 'the garment of your weaving'. Apart from the presence of the deictic particle (replacing ana if this would otherwise be required), the relative clause is identical to a main clause: Maori permits the omission of subject NPs and prepositional phrases (and indeed Maori omits prepositions in relative clauses when they would govern a relative pronoun in English). Consequently this interpretation of relative clauses as referring to the embodied fact rather than to the antecedent, and having an operand rather than appositional relationship, would seem a straightforward matter.

A similar argument may be put forward for the Po/AN language Japanese. Here the relative clause precedes its 'antecedent':

tegami o kaite iru hito
letter Acc writing is person

'(the) person who is writing a/the letter'

But the relative clause can in fact stand on its own as a main clause - at least in the non-honorific register:

tegami o kaite iru
'he is writing a latter'
(or in fact any subject pronoun)

It may be noted that a large class of adjectives in Japanese are in fact intransitive verbs and when such 'adjectives' are used attributively, they make use of a relative clause:

kono hon wa akai 'this book is red'
this book Top is-red

akai hon 'red book'
is-red book

If we place koto, (abstract) thing, fact, after akai to give

akai koto

the phrase does not mean a 'red fact' but the fact of something being red. This leads us to a general analysis of prenominal clauses, of the form:

[[S] N]

with the meaning 'N pertaining to S', rather than 'N being S'.

7.5 Adjunction

In my last three sections I considered adjectives as operators, appositional items, and as operands, and from the positive evidence for such structures showed that variation in analysis was to be found both between and within languages: thus it appeared in English that little was gained from a uniform treatment of adjectives. In this section I shall show that a fourth form of analysis is also necessary, which I shall term adjunction. I shall additionally show that this perhaps provides the greatest insights into the study of adjectives.

In section 7.2, I suggested that only adjectives which could not occur predicatively should be treated as operators. Difficulties occur however with adjectives such as apparent and strong in:

His guilt was apparent.
Their support was strong.

Although the first adjective appears to differ from the second in having a modal content, their treatment is in fact very similar. There is no difficulty about treating them as predicative, but if we take expressions such as

apparent culprit
strong supporter

we have no corresponding predicative forms:

* That culprit is apparent,
* That supporter is strong

unless the latter refers to physical strength.

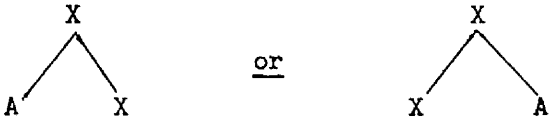
We might regard this as an intermediate situation between operator adjectives and appositional adjectives. Thus apparent is classifiable as D(i), or Modal, and strong as E, or Dimensional. On the other hand we find no difficulty about F, Value:

That batting is lousy.
That batsman is lousy.

But what seems to be critical is that these partially predicative adjectives may only be predicative when a direct description is involved, as when we describe somebody's guilt as apparent. Thus in 'apparent culprit' it is the guilt, not the culprit, that is apparent.

To deal with this problem I propose a relationship of adjunction, whereby the adjunct directly describes the adjoinee, or else describes the most appropriate concept implicit within the adjoinee expression. A critical test for adjective being used in both manners will be that only in the second case it may be used predicatively. For this scheme to work, it must be noted that value adjectives such as lousy must refer directly both to persons and qualities or activities associated with them. We must therefore treat lousy batsman as lousy for a batsman in the same way as we treat small elephant as 'small for an elephant, rather than compare it with strong supporter and apparent culprit.

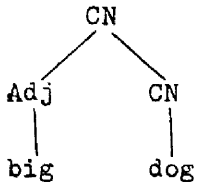
To depict the relationship I shall represent it in the form



where A is the adjunct. No attempt will be made either to express the X as X/A or the A as X/X. It is perceivable in the above syntactic structures, in the present work at least, by simply being there. In the linear form, however, I shall represent it by

$$[A+ \quad X] \quad \text{or} \quad [X \quad +A]$$

We would thus have

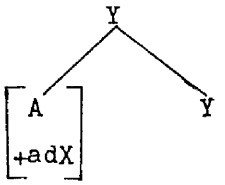


or simply [big+ dog]

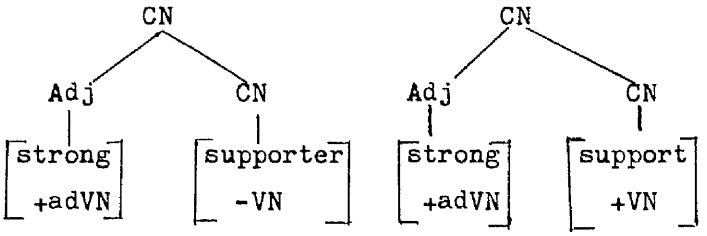
where for the rest of this work I shall follow Montague analyses in using CN to denote a common noun expression, be it a simple noun or larger expression.

I shall now formally consider the question of the semantic content of the adjunctive structure. The proposal I shall put forward is as follows:

If an adjunct A is semantically ad-X, ie. it most naturally describes an X, and takes part in a construction of the form



then the adjunct will describe the most accessible [+X] item within X, which may of course be the adjoinee Y itself. Thus we may have, as two examples,



where VN stands for 'verbal noun'.

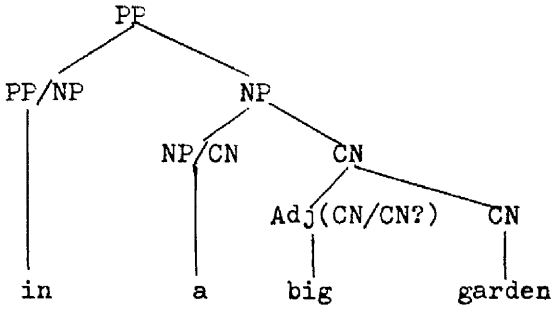
Thus in 'strong support' we have a situation where strong is [+adVN] and support is [+VN], so that here 'strong' not only adjoins its noun, but may also directly describe it. This situation does not prevail for 'strong supporter', where the noun is [-VN]: consequently the adjective strong must describe the most accessible [+VN] concept, i.e 'support'.

What is the evidence for this adjunctive approach?

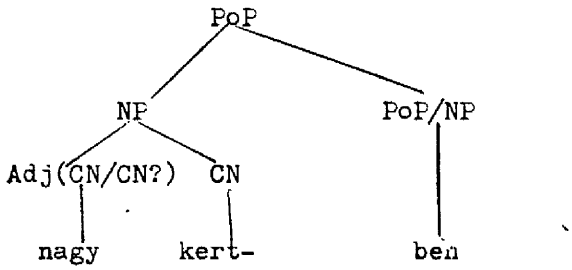
The first evidence I shall take is from Hungarian, from the phrase

nagy	kert-	ben	'in a big garden'
big	garden	in	

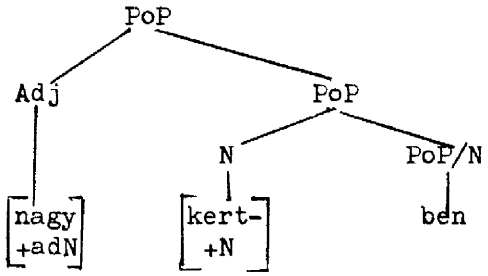
where in English we might have



But in Hungarian the postposition is attached to the last item of the NP. We may either make the word kertben straddle the phrase boundary:

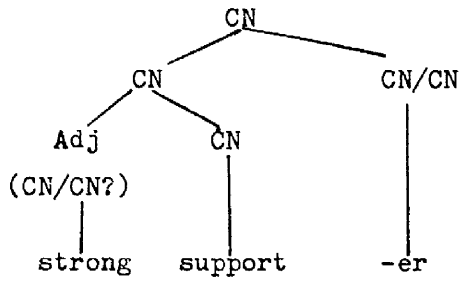


or simply adjoin the nagy to the postpositional phrase:



Note here that nagy has the feature [+adN] and kert- the feature [+N], so that nagy may describe kert-. Note also that the postposition in the second account has become PoP/N rather than PoP/NP, since it only requires a noun, not even a common noun expression (CN), to make up a postpositional phrase (PoP). It also solves for Hungarian but not English the problem of having an NP node dominating a common noun expression (CN) without a determiner expression (NP/CN) being present.

If such a device exists in one situation, what remains to be argued is whether we are constrained syntactically as to what such an adjunct describes. This may of course vary from language to language. But in English we would be unlikely to want an analysis such as



where 'strong support-' is a constituent, at least not unless we permitted words to straddle phrase boundaries.

On the other hand, it seems unlikely that we would wish to have two separate lexical specifications for strong, one for 'support', directly describing it, and another for 'supporter', describing something within the noun. We could of course define strong in vague

terms such as 'showing strength'. But then it is perfectly acceptable to say that our supporters show strength, so that we cannot thereby bar assertions that our supporters are strong, in a situation when we have strong supporters. If however we permit a strategy of describing the most accessible relevant concept within the adjoinee, whilst only permitting direct description for the predicate expressions, since adjectives are not then adjuncts to the subject, we now have something that will fit the bill.

The proposed properties of adjuncts don't merely 'fit the bill' for an intensifying adjective such as strong in strong supporter : it also clearly crops up with other adjectives referred to above which have been thought to be straightforward modal or privative intensional adjectives. Even here we may find the predicative/non-predicative pairs:

His crime was apparent,
His guilt was alleged (by someone),
His guilt was probable,

as opposed to

* The criminal was apparent,
* The culprit was alleged (by someone),
* The culprit was probable,

though we may have the attributive forms: 'the apparent criminal', 'the alleged culprit', 'the probable culprit'.

The second piece of evidence I wish to put forward is the fact that there is a very close correspondence between some adjectives and related adverbs. We have already seen this in section 7.2 for former and formerly, and also for apparent and apparently. We also find this for certain intensifier adverbs such as extreme and extremely:

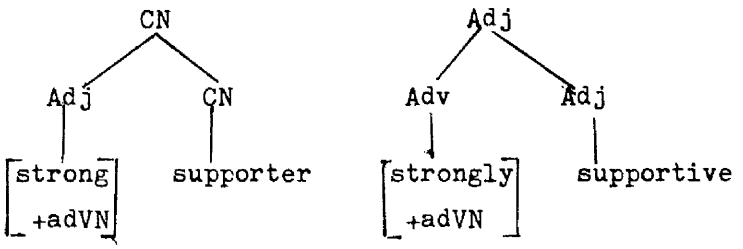
extremely clever person,
extreme cleverness.

We have here a parallel in meaning between adjective and related adverb, but if we look at another example:

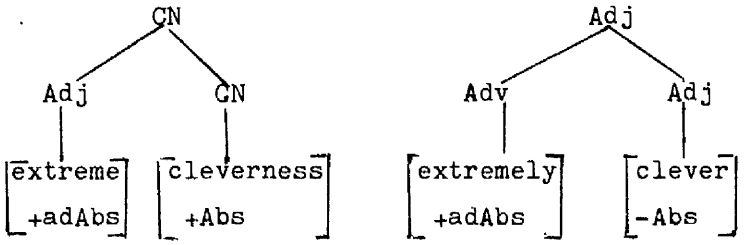
strongly supportive,
strong supporter,

we also find that the semantic relationships between adjective and

noun, and between adverb and adjective, are identical. However no problems arise if we treat strongly as an adjunct:



Similar structures may also hold for intensifying adjectives and adverbs:



though in this case, 'extreme' may directly describe 'cleverness', since it has the feature [+Abs] for an abstract noun.

A third piece of evidence comes, once again, from Tagalog, where it would appear that the particle na/ng can link an adjunct to its adjoinee in either order:

A na X or X na A.

I have already illustrated this construction for adjective and genitival expressions. The relevant evidence for the semantic properties of adjuncts occurs with a limited number of adjectives which may also act as intensifiers of other adjectives. These are[6]

lubha,	serious	masyado,	excessive
totoo,	true	tunay,	true
talaga,	true		

These may combine with another adjective in the same way as an adjective might combine with a noun:

Totoo	-ng	pagod	ang	kabayo
true	+	tired	Top	horse

'the horse is very tired'

or Pagod na totoo ang kabayo.
+

Thus the above expression does not mean that the horse is both true

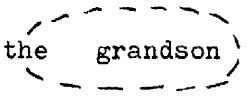
and tired, but that it is truly tired. Thus the adjective totoo describes not the tired creature, but the quality of tiredness.

In this chapter I have considered the relationship of adjectives and adverbs to the items they adjoin, and argued for a structure either additional to or defined separately from operator-operand.
Much of my discussion involved the question of whether adjectives and adverbs should be treated as a separate semantic structure, or simply one which was indifferent as to what constituted the operator and what constituted the operand. Or perhaps the choices are not incompatible.

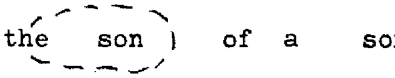
One answer is certainly that for the purposes of cross-language comparison a separate concept is necessary for the purpose of explaining the statistical patterns. It would then appear that there may in fact not be one single argument for treating adjuncts variously as operators or operands, but a different argument for each language: in English it seemed best to treat some adjectives as operators and some as adjuncts.

However I would now like to look at two further arguments which further illustrate the ambivalence of the situation, so that perhaps after all the choices are not in fact incompatible. The first arises when semantic regions are involved, and the second involves the identification, in operator-operand expressions, of which is the operator and which is the operand.

Turning to semantic regions, we may recall expressions such as



which contrasted with



where the definite determiner indicates that the addition of the operator standing within its region, be it grandson, son or the like, yields a unique reference. But we may also envisage a situation in which the addition of an adjunct also yields a unique reference, ie. a language in which 'the white billiard-ball' is treated as something like

the white one of the billiard-balls .

Similarly it may be possible to argue that the addition of head or adjoinee expression may also create a unique expression:

the old billiard-ball which is white .

A good example of the like treatment of operators, adjuncts and even adjoinee expressions as semantic regions may be illustrated in Tagalog. The English expression

the pencil of the child

may be translated as follows:

ang	lapis	na	sa	bata
Top	pencil	+		child

Here the ang contrasts with sa and ng (see below for latter) in topicality, while the latter two contrast with each other in grammatical function, the sa including indirect object and genitive among its functions. The ang also contrasts with them in requiring that its phrase must be definite. Thus we can say that when lapis, pencil, is added to sa bata, of the/a child, a unique reference is obtained, something which does not necessarily occur with ng and sa.

Other evidence that ang, ng and sa are second-order operators, ie. surround semantic regions, arises when we consider prepositions. These do not replace the above items but rather occur before either ng or sa[7]:

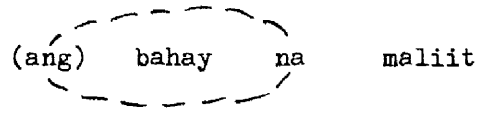
Bumili	ako	ng	damit	para sa	bata
bought	I (Top)		clothes	for	child

'I bought some clothes for the child'

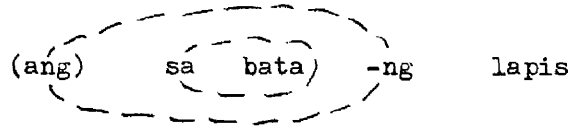
The above construction may also be applied to adjectival modifiers:

(ang) maliit na bahay
small

or alternatively



Note too that the possessive phrase may also precede:



so that a genitival expression commencing with sa is just as much an adjunctive expression as maliit, small.

Note that according to the above analysis nesting of regions is permissible. We should also observe that that form -ng attaches itself to the last items of the region it terminates - when this ends with a vowel - even when this is not a head expression. Precisely how the na/ -ng should be linked up, both semantically and syntactically, is not clear. If however it is a second-order operator, it gives rise to very interesting and revealing implications for phrase-structure grammar, if this is to reflect the above relationships.

A final point may now be discussed: can we in any way avoid the circularity that has been reverted to once Vennemann's principle of category constancy has been dropped? In particular, can we show what adjectival operators have in common with other operators, such as prepositions, acting as operators on non-subject NPs? I believe that a very simple principle may be adhered to:

> If a phrase permits the addition of adjuncts,
> it must either (a) be itself an adjunct, or (b)
> act as an operand.

What this means for example is that if we take the adjective red-haired as an adjunct to typist in the CNP expression 'red-haired typist', the underlined items will be operators in the following examples:

> This ad will be of interest to red-haired typists,
> We do not employ red-haired typists.

unless of course we have reason to treat the object expression as an adjunct in the above cases. Further examples, using operator adjectives are as follows:

She is our chief red-haired typist.

She is a former red-haired typist.

These examples are interesting in that they bring out rather sharply the way chief and former interact with the expression following them.

These examples are very interesting in that they suggest that there is a similarity in the way certain items, such as prepositions, transitive verbs and adjectives such as 'chief' and 'former' interact with expressions following them. This line of thought is also followed in Siegel (1980), who contrasts the behaviour of such adjectives with 'measure' adjectives. Thus the sentence

She is a tall red-haired typist

does not necessarily mean that hair-colour is relevant to the yardstick by which a typist is considered to be tall. Siegel is of course primarily concerned with the analysis of such measure adjectives, but the example is also of interest with regard to the question of adjuncts and their role in identifying operands.

Given the semantic analysis of adjuncts, and given the above principle for identifying operators and operands in an operator-operand structure, we have a very clear hypothesis to test: it must be noted that such further testing is indeed necessary, since one thing has simply been seized upon, which NPs after transitive verbs and prepositions, and CNPs after adjectives such as main and former have in common. This leads us to my last two chapters, of which chapter 8 will spell out the implications of the above findings for verbs, in particular the interaction of finite verbs with subject and object, and chapter 9 will draw together the semantic and statistical implications of this work, in particular its consequences for language change.

8. Verbs

In this study I have sought evidence primarily from the study of verbs and adjectives: the last three chapters have in fact explicitly considered the problems of adjectives, and it was as a result of this investigation that the semantic concept of adjunction was evolved, and the finding that if other constructions, such as PPs and VPs involve operator-operand structures, it is not the NP that acts as operator, as Bartsch and Vennemann argue, but the items governing them.

The analysis of verbal constructions has also featured throughout this study, but in contrast to the analysis of adjectives, it has only featured implicitly, in particular as part of stratal and prediction chain statistics. Thus a number of strands have emerged which I shall attempt to bring together in this brief chapter. One of the most important observations thrown up in this work is that we cannot necessarily regard tense as a first-order operator on propositions or clauses: for English at least we must reject a semantic structure such as

[Tense [Clause]] [Predicate]] . [Subject]] .

This is not because English differs from many VSO or VOS languages which do in fact place tense marking at the start of a finite clause, but because of examples cited in chapter 1 (§1.3), where the tense of the verb depends upon the time of the subject, as in '1977 prices', or at least the time when the predication applied. Thus we must regard tense instead as a second-order operator, reflecting the subject-predicate breakdown, giving something like

[Subject Tense [Predicate]] .

or [Predicate Tense [Subject]] .

This question must particularly be borne in mind when we consider SVOPr languages, where SV is relatively inconsistent with Pr. If operators consistently precede operands in Pr languages, it would appear for these languages at least, that we cannot treat the subject NP as an operand to the verb.

Turning from the subject-predicate relationship, we may also consider the verb-object relationship: the statistical data showed a very strong correlation between VO/OV and Pr/Po, not only in synchronic terms, but also in terms of the stratal analysis:

$$\underset{(9)}{\text{VO/OV}} \xleftrightarrow{7} \underset{(7)}{\text{Pr/Po}}$$

Thus all the language families in the sample showing Pr/Po variation also showed a correlation between Pr/Po and VO/OV. A possible explanation involving the operator-operand structure would treat NPs as operands in each case:

$$\begin{bmatrix} \text{Verb} & [\text{NP}] \end{bmatrix},$$
$$\begin{bmatrix} \text{Preposition} & [\text{NP}] \end{bmatrix}.$$

in effect the mirror-image of the Natural Serialisation Principle where it is the NPs that act as operators.

In examining the above questions we must note that whilst the behaviour of tense as a second-order operator represents a strong constraint on how we analyse finite clauses, there are a number of options open to us in relating subject and object to the verb. Before we go into this however, it is important to note a mathematical process of operator composition. We may see such composition in the concept 'grandfather of', which is composed of 'father of' and 'parent of' in that order:

$$\begin{bmatrix} \text{grandfather of} & [\text{X}] \end{bmatrix}$$
$$\begin{bmatrix} \text{father of} & [\text{parent of} & [\text{X}]] \end{bmatrix}$$

In this example we are of course not exemplifying operators and their composition on a syntactic level, so much as relating the meanings of lexical items, and the composition of the above in the opposite order gives us 'paternal grandparent'. In general I shall represent the composition of P and Q by P*Q, so that

$$[\text{P*Q}[\text{X}] = [\text{P}[\text{Q}[\text{X}]]].$$

As mentioned before, there are a number of ways of relating subject and object to the verb. Now the principle enunciated at the end of the last chapter precluded any NPs from acting as operators if they included adjuncts: they themselves could be adjuncts or act as

operands, but they could not act as operators. But this does not preclude pronominal subjects - or indeed pronominal objects - from acting as operators:

? [We [love [Jack]]]

Can we do likewise with French? The above example translates into French as

? [Nous [aimons [Jacques]]]

In fact, since French conjunctive pronouns are closely attached to their verbs, it might be better to represent the above expression as

[Nous * aimons [Jacques]].

We may further note that an object pronoun precedes its verb:

[nous * l' * aimons],

though there is nothing to stop us in the absence of a following nominal object from writing:

[nous [l' [aimons]]].

But what happens if we say

Nous l'aimons Jacques,

or Nous l'aimons lui ?

Here the 'Jacques' and the 'lui' amplify the object pronoun. We now appear to have two alternatives:

[nous * l' * aimons [Jacques]]

or else

[[[nous [l' [aimons]]]] +[Jacques]]

where 'Jacques' is treated in the latter as an adjunct: either it directly describes 'nous l'aimons' (which it doesn't), or it describes a relevant concept within that expression, ie. the object pronoun l(e). We must bear in mind though that the above data on its own is insufficient to determine which structure is appropriate.

I shall now turn to consider the same sentence in Spanish:

(Lo) amamos a Diego.

This objective 'a' is used with definite human objects, with vacillating usage before indefinite human objects, and before non-human

animate objects. It is also used with verbs such as preceder, precede, where context cannot distinguish inanimate subjects and objects. We would appear to have a similar choice to that of French:

? [lo * amamos [$_{pp}$a [$_{NP}$ Diego]]],

and ? [[lo [amamos]] +[a [Diego]]].

- where the 'lo' may of course be omitted, or indeed be replaced by the Dative pronoun le.

The latter analysis would appear to be more attractive, in that 'a Diego' is not closely attached to the verbal complex, so that, for example, a VSO ordering is perfectly permissible. We might therefore propose that in French the post-verbal object NP is an operand, whilst the object NP or PP in Spanish is an adjunct, which has the ability to adjoin itself either before or after the verb. The evidence for each approach requires careful examination. It may be for example, that when a double negation occurs, as in the Spanish

No veo a nadie. 'I don't see anybody'

that 'nadie' does not cancel out the first negative, 'no', since it merely adjoins the negative expression and amplifies it:

[[No veo] +[a [nadie]]]

(where the 'no veo' is left unanalysed here)

It may be possible to establish that adjunction is a necessary condition for non-cancellation. If this is correct, and if the post-verbal object NP in French is an operand in expressions such as

Je (ne) vois personne

then it would follow that the 'ne', which may be dropped in colloquial speech, has lost its negative content.

In the above discussion I have shown firstly that nominal object NPs need not be operands to their transitive verbs: they may also be adjuncts, perhaps identifiable by phenomena such as double negation. Secondly it appears that the pronominal subject and object may in fact be anything. We may thus see from the evidence that even though my initial argument refuting the Bartsch/Vennemann scheme

[[Verb] NP]
[[P] NP]

leads us to switch operator and operand, examination of the precise construction may lead us to prefer something which involves a more flexible adjunct construction, which may amongst other things permit non-dominant orderings:

Al silencio siguió el bullicio 'Tumult succeeded the silence'

A very strong argument may thus be put forward in languages with object marking, for treating object phrases as adjuncts rather than as operands. Thus as well as having the pattern

$$[\,V\quad[NP]\,]$$

we may also have

$$[\,[V]\;+[\,Obj\;[NP]\,]\,].$$

If such a situation combines with one in which subject phrases too are treated as adjuncts, permitting both

$$[\,[V]\;+[\,Subj\;[NP]\,]\;+[\,Obj\;[NP]\,]\,]$$

and

$$[\,[V]\;+[\,Obj\;[NP]\,]\;+[\,Subj\;[NP]\,]\,],$$

then the concept of a VP becomes much more questionable, in finite clauses at least.

I would now like to look more closely at the relationship between subject and verb, and since the relevant statistics, as outlined in chapters 3 and 4, appear to be quite complex, so we must expect the analysis to be quite complex. I have already suggested that we require the intervention of tense as a second-order operator, and this permits structures such as

$$[\,(\text{Predicate}\quad\text{Tense})\quad[\,\text{Subject}\,]\,]$$

Here the predicate acts as a first-order operator on the subject and the tense acts as a second-order operator on the predicate.

In languages such as Spanish, where the verbal complex can stand on its own, we may analyse 'hablábamos', we were speaking, as

$$[\,(\text{habl\'a-}\quad\text{ba-})\quad[\text{mos}]\,]$$

speak Imperf 1pl

Any item following it and perhaps before it as well may then be treated as adjuncts to the finite verb.

In the above analysis I have pointed out how my identification of operators and adjuncts can be extended without difficulty to the analysis of finite clauses with a clearly identifiable verb-complex. We may now consider how we should treat languages such as English, where the structure

$$[\ NP \quad VP\ ]$$

has greater justification. Since an NP can admit adjuncts we cannot according to my analysis treat the subject NP simply as an operator:

$$[\ NP \quad [\ VP\ ]\ ].$$

All we know for the moment, as a result of the analysis of tense, is that second-order operators seem to be involved in some way. We may also note however from the prediction chain analysis that for non-inflected subjects, VS/SV was more closely correlated with NG/GN than with Pr/Po, at least among Pr languages, and it may well be the case that among these languages SV and GN employ similar semantic structures. We may note that the 's of English may be considered to be a second-order operator, or part of one:

? My uncle 's idea of a night out .

Here we have a very similar problem to that encountered with the subject NP: we cannot treat the genitive NP simply as an operator since it permits adjuncts. Nevertheless the NP plus the 's acts as a definite determiner. Thus the interaction of NP's with second-order operators requires further investigation, particularly whether the dashed lines above are marked out appropriately.

I have now considered the inconsistency of SV and Pr in SVOPr languages. I shall next turn to look at postpositional SVO languages - SVOPo - and the situation in which subject position shows greater consistency with postpositions than does object position.

The hypothesis that operands consistently precede operators in Po languages would permit the following structure:

$$[\ [\ [\ NP\ ]\ \ V\ ]\ \ (+)[\ NP\ ]\ ]$$

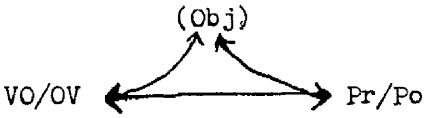
(S) (V) (O)

The relationship here of the object to the verb raises similar questions to that of the subject in Pr languages. What I wish to

concentrate upon though is whether we have any explanation of the phenomenon I observed in chapter 3 - that such languages appear to lack passives. In this regard we must look not so much at the relationship between the object and the verb, but at the status of subject and verb within the clause: the above structure suggests that we have a kind of verb-subject phrase or VSP, rather than the normal VP combining verb and object. But what is interesting about such a VSP phrase is that its semantic content corresponds to that of a different kind of 'normal' VP: not one combining verb and object, but one combining passive verb and agent. In this case what we have is not so much a motivation for the lack of passives, as the lack of motivation for such passives, since the 'active' form already provides a phrasal combination of verb and agent.

This hypothesis about passives is undoubtedly an interesting one, and there was additionally some indication in chapter 3 that when languages were uncertain in their pre/postpositionality, there was also uncertainty about some of the verbal constructions and their status as passives. However the investigation of the question in this present study can only be limited: we cannot for example say why passives should not occur side by side with syntactic active constructions, even if their semantic structure is similar. What is of interest here though is that we have a situation in which there is something more than just a violation of a statistical pattern, and further that the identification of operands and adjuncts is not purely a circular exercise.

This short chapter has in effect taken as a central assumption that tense is a second-order operator: as a result it provides an ideal account for languages in which the verb complex is critical to finite clauses, and where subject and object NPs can be added on as adjuncts to that verb complex, amplifying a subject or object appearing explicitly or implicitly within the verb complex. However in so doing we open up the possibility that statistics on VO/OV ordering conceal more than one type of relationship, ie. one where the verb is an operand, and the other where the verb is an adjunct. In this we may ask why there is such a high overall correlation between VO/OV and Pr/Po? Surely there should be no correlation between an adjunct object and Pr/Po? The answer lies in the split chain drawn up in chapter 4. The chain relevant to object position is as follows:



Now what this shows us is that when there is no object marking, we can indeed correlate VO/OV with Pr/Po: such language may well treat both direct and pre/postpositional objects as operands. On the other hand, when case-marking does occur, VO/OV correlates only with the direct-object marking. In other words the placing of an adjunct phrase is influenced by the ordering within it.

The above example brings out an interesting application of the prediction chain: the dominant ordering of an adjunctive object phrase correlates with the position of object-marking within the phrase, which in turn correlates with Pr/Po, perhaps because object-marking and pre/postpositions are operators. As a result the correlation between VO/OV and Pr/Po is only indirect in such examples, but two types of relationship have been isolated: that between verb and object phrase, and that between the object-marking and what it marks.

9. Identification of Constraints

0.

In the past four chapters (chs. 5-8) I have looked more closely at problems pertaining to semantic questions and argued that at the very least we must separate out operator-operand structures from those of adjunct and adjoinee. In this last chapter I wish to tackle two remaining problems. Firstly I shall consider the consequences of trying to fit the results of the semantic investigations into the statistical framework: in particular I shall show that a prediction chain account is not appropriate for the adjectival hierarchy considered in chapter 6. Apart from the fact that we do not have enough data on rare combinations such as Nsm/NumN to order the items on the hierarchy relative to each other, it also becomes clear that we must take into account the central influence of Pr/Po, leading to a pattern of closed chains.

Secondly I shall look further at the question of semantic explanations, and argue that the distinction between first- and second-order operators, in particular the isolation of definite determiners, is essential for the explanation of certain asymmetric statistical phenomena in the data. A similar approach may be adopted for indefinite determiners, but I shall show that on both semantic and statistical grounds, we may treat indefinite determiners as sharing the properties of both first- and second-order operators.

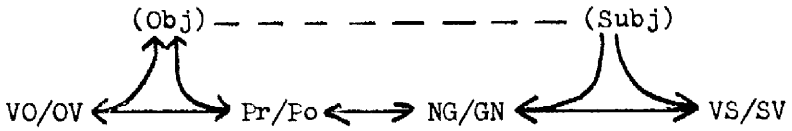
I shall conclude this chapter with a summary of the statistical patterns I have identified in this work, and the semantic constraints that appear to motivate them.

9.1 Limitations on the Prediction Chain

It will be recalled from chapter 3 that a prediction chain was devised from Hawkins' data in the form

$$\text{VSO/SVO/SOV} \longleftrightarrow \text{Pr/Po} \longleftrightarrow \text{NG/GN} \longleftrightarrow \text{NA/AN.}$$

In choosing the data from my own sample, modifications to the prediction chain seemed necessary, and in chapter 4, which looked at case-marking, a revised form emerged:



where in practice 'Obj' and 'Subj' are in fact difficult to identify as separate factors. The status of NA/AN has to be considered further.

As observed in chapters 3 and 4, the position of VS/SV is very interesting: when there is no case-marking for subject, the link

$$\text{Pr/Po} \longleftrightarrow \text{NG/GN} \longleftrightarrow \text{VS/SV}$$

is by-passed, though more evidence would be desirable.

The task of this section, including the checking of NA/AN, will be to consider how we can extend the prediction chain by means of the data on adjectival position obtained in chapter 6. The question arises of whether we can except a chain of the form:

$$\text{AllN/Nall} \longleftrightarrow \text{NumN/NNum} \longleftrightarrow \text{smN/Nsm} \leftrightsquigarrow \text{NA/AN,}$$

where the above items observe the order of the hierarchy discussed in chapter 6.

Note the form

$$\text{smN/Nsm} \leftrightsquigarrow \text{NA/AN,}$$

is equivalent to

$$\text{smN/Nsm} \longleftrightarrow \text{AN/NA,}$$

but permits us to build up a circular chain pattern, in which the Möbius effect referred to in chapter 1 may be represented.

The critical problem in evaluating such a chain is the dearth of languages with AN/Nall, AN/NNum and AN/Nsm ordering, but other

difficulties arise: we are in effect conflating two different patterns of behaviour, one for Pr languages, approximating to

$$\text{Pr} \supset (\text{Nall} \Rightarrow \text{NNum}),$$
$$\text{Pr} \supset (\text{NNum} \Rightarrow \text{Nsm}),$$
$$\text{Pr} \supset (\text{Nsm} \Rightarrow \text{NA}).$$

at least when the correlation is strong[1], and one for Po, approximating to

$$\text{Po} \supset (\text{allN} \Rightarrow \text{NumN}),$$
$$\text{Po} \supset (\text{NumN} \Rightarrow \text{smN}),$$
$$\text{Po} \supset (\text{smN} \Rightarrow \text{AN}).$$

For this reason it is not particularly revealing to make use of a prediction chain here, since it merely blurs information that we have already captured.

Table 9.1a Prepositional Languages

	allN	Nall
NumN	Cairene	Cemuhi
	Chamorro	Fijian
	Greek	Maori
	Palauan	Niue
	Tagalog	YUROK
	Tsou	(5)
	Vietnamese	
	Welsh	
	Yapese	
	CAYUVAVA	
	JACALTEC	
	SHUSWAP	
	(12)	
NNum	Fulani	Sango
	Hausa	Thai
	Iraqw	(2)
	Margi	
	Tiv	
	Yoruba	
	(6)	

Not listed above:

Position of 'all' varying:

Berber, Kinyarwanda, Mandarin, Mâsai, Zulu

Position of numerals varying:

Malay, SULKA

No data for position of 'all':

Maung, Tiwi, ZAPOTEC, OLO.

(Total Pr languages 36.)

Table 9.1b Postpositional Languages

	allN	Nall
NumN	Finnish	Burushaski
	Hindi	(1)
	Kannada	
	Newari	
	Turkish	
	CARIB	
	LUISENC	
	PIRO	
	QUECHUA	
	(9)	
NNum	Kanuri	Bawm
	(1)	Burmese
		Lugbara
		Mandingo
		Nubian
		Wati
		DAGA
		KATE
		NASIOI
		SENTANI
		SIROI
		WASKIA
		(12)

Not listed above:

Position of 'all' varying: Nama

Position of 'all' infixed: Ewe

No data for position of 'all':

FORE, GUARANI, GUAYMI, HIDATSA, MARIND, POMO(E), ZOQUE.

Languages varying NumN/NNum: Abkhaz, Ijo, Japanese

Data unclear for NumN/NNum:

Bardi, Djingilo, Ngandi, MAIDU.

(Total Po languages 39.)

Table 9.2a Prepositional Languages

	smN	Nsm
NumN	Cemuhi	Berber
	Chamorro	Cairene
	Greek	Fijian
	Mandarin	Maori
	Maung	Niue
	Tiwi	Vietnamese
	Tsou	Welsh
	Yapese	JACALTEC
	CAYUVAVA	ZAPOTEC
	SHUSWAP	(9)
	YUROK	
	(11)	
NNum	Tiv	Fulani
	(1)	Iraqw
		Kinyarwanda
		Margi
		Thai
		Yoruba
		Zulu
		OLO
		(8)

Not listed above:

Languages varying position of 'small':

Hausa, Masai, Palauan, Sango, Tagalog

Position of numerals varying:

Malay, SULKA

(Total Pr languages 36.)

Table 9.2b Postpositional Languages

	smN	Nsm	
NumN	Burushaski	GUARANI	
	Finnish	(1)	
	Hindi		
	Kannada		
	Nama		
	Newari		
	Turkish		
	QUECHUA		
	ZOQUE		
	(9)		
NNum	MARIND	Bawm	POMO(E)
	(1)	Burmese	DAGA
		Ewe	KATE
		Kanuri	NASIOI
		Lugbara	SENTANI
		Mandingo	SIROI
		Nubian	WASKIA
		Wati	(17)
		GUAYMI	
		HIDATSA	

Not listed above:

Languages varying smN/Nsm:

CARIB, LUISENO, PIRO

Languages varying NumN/NNum:

Abkhaz, Ijo, Japanese, FORE

Data unclear for NumN/NNum:

Bardi, Djingili, Ngandi, MAIDU

(Total Po languages 39.)

Table 9.3a Prepositional Languages

	smN	Nsm
NA	Cemuhi	Berber
	Tiv	Cairene
	Yapese	Fijian
	SULKA	Fulani
	(4)	Iraqw
		Kinyarwanda
		Malay
		Maori
		Margi
		Niue
		Thai
		Vietnamese
		Welsh
		Yoruba
		Zulu
		JACALTEC
		ZAPOTEC
		OLO
		(18)
AN	Chamorro	-
	Greek	
	Mandarin	
	Maung	
	Tiwi	
	CAYUVAVA	
	SHUSWAP	
	YUROK	
	(8)	

Languages not listed above:

Varying NA/AN: Palauan, Sango, Tagalog, Tsou

Varying smN/Nsm: Hausa, Mâsai

(Total Pr languages 36.)

Table 9.3b Postpositional Languages

	smN		Nsm
NA	-		Abkham
			Bawn
			Burmese
			Djingili
			Ewe
AN	Bardi	(NA)	Kanuri
	Burushaski		Lugbara
	Finnish		Mandingo
	Hindi		Nubian
	Ijo		Wati
	Japanese		GUARANI
	Kannada		GUAYMI
	Nama		HIDATSA
	Newari		POMO (E)
	Turkish		DAGA
	MAIDU		KATE
	QUECHUA		NASIOI
	ZOQUE		SENTANI
	FORE		SIROI
	MARIND		WASKIA
	(15)		(20)
		(AN)	-

Not listed above:

Languages varying NA/AN:

Ngandi, LUISENO, PIRO

Languages varying smN/Nsm:

CARIB

(Total Po languages 39.)

It is however revealing to take correlations separately for pre- and postpositional languages.

The basic data is given in Tables 9.1-3. If we look first at Pr languages, we may summarise the statistical patterns in Table 9.1 below:

Table 9.4

(Prepositional)

	allN	Nall	smN	Nsm
NumN	12(12.2)	5(4.8)	11(8.3)	9(11.7)
NNum	6(5.8)	2(2.2)	1(3.7)	8(5.3)

	smN	Nsm
NA	4(9.8)	18(12.2)
AN	8(3.2)	0(4.8)

The above data is taken from the parts a of Tables 9.1-3. We may first of all observe that there is no discernible correlation between allN/Nall and NumN/NNum: instead we see that roughly 72% of the Pr languages are allN and 68% are NumN, with little correlation between the two. Lack of correlation has two consequences: firstly it is not only inappropriate to employ the implication

$$Pr \supset (Nall \dot{\supset} NNum)$$

it is simply incorrect, in that according to the above table, approximately 68% of Pr/Nall languages also have NumN ordering (though note the discussion of Fiji, Maori and Niue in section 6.2.1 of chapter 6). We may secondly note that if the direct correlation is so weak, there is nothing to be explained by such a link in the prediction chain.

We may note however that a correlation does occur between the next pair, NumN/NNum and smN/Nsm. Here we are justified in putting forward the implication

$$Pr \supset (NNum \dot{\supset} Nsm),$$

since we have only one example of an exception to this implication, the language Tiv. The same may be said even more strongly of the implication

$$Pr \supset (Nsm \dot{\supset} NA),$$

since there are no exceptions in the sample to this implication, and Table 9.4 shows that the correlation is also stronger, since a random

distribution would give 4.8 Nsm/AN languages, instead of none.

We must further note however that the correlation between NumN/NNum and smN/Nsm could well be due to their respective correlation with NA/AN: in such a case we would have a prediction chain with NA/AN in the middle, but it is not possible to test this properly, since there is little data on AN/NNum and AN/Nsm languages.

If we turn to postpositional languages, for which the data is summarised in Table 9.5, we find some strong differences in the pattern:

Table 9.5

(Postpositional)

	allN	Nall	smN	Nsm
NumN	9	1	9	1
NNum	1	12	1	17

	smN	Nsm
NA	0	20
AN	15	0

The above data is derived from parts b of Tables 9.1-3 and here we find very strong correlations in each case, providing further evidence that an attempt to use a prediction chain to express the adjectival hierarchy would be problematic, in that it would conflate two different groups of languages - one Pr and one Po - with different statistical patterns.

The above analysis has been carried on a separate basis for Pr and Po languages, and served its purpose in arguing against the use of a prediction chain for items on the adjectival hierarchy. The use of such separate analyses is limited however, since items such as numeral position and the position of 'all' show great stability within language families, even when these show variation for Pr/Po. I must therefore use other methods to consider the important question of whether adjacent items on the hierarchy have a direct influence upon each other, or whether instead their apparent correlation represents a similarity in their separate correlations with Pr/Po. To deal with this I would first like to carry out a stratal analysis of the supposed link between position of 'all' and the position of numerals:

allN/Nall $\longleftrightarrow$ NumN/NNum.

The relevant data is detailed in Table 9.6. Languages with more than one ordering are entered for each ordering and indicated by =.

Table 9.6a Stratal Analysis

	NumN		NNum	
	allN	Nall	allN	Nall
Austronesian (10)	Malay= Chamorro Palauan Tagalog Tsou Yapese	Cemuhi Fijian Maori Niue	Malay=	-
Niger-Kordofanian (9)		Ijo=	Kinyarwanda= Zulu= Fulani Tiv Yoruba	Kinyarwanda= Zulu= Ijo= Mandingo Sango
Australian (6)	?	?	?	Wati
	(note that Maung, Tiwi have NumN)			
Afroasiatic (5)	Berber= Cairene	Berber=	Hausa Iraqw Margi	-
Nilo-Saharan (4)	-	-	Mâsai= Kanuri	Mâsai= Lugbara Nubian
Sino-Tibetan (4)	Mandarin= Newari	Mandarin=	-	Bawm Burmese
Ural-Altaic (3)	Japanese= Finnish Turkish	Japanese=	Japanese=	Japanese=

(continued on next page)

Table 9.6a Stratal Analysis (continued)

	NumN allN	NumN Nall	NNum allN	NNum Nall
Indo-European (3)	Greek Hindi Welsh	-	-	-
other (6)	Nama= Vietnamese Kannada	Nama= Abkhaz= Burushaski	-	Abkhaz= Thai

Not listed for infixation of 'all':

Ewe (Niger-Kordofanian)

Not listed for lack of data:

Bardi, Djingili, Maung, Ngandi, Tiwi (Australian)

Table 9.6b. Stratal Analysis

	NumN		NNum	
	allN	Nall	allN	Nall
Andean-Equatorial (4)	CAYUVAVA PIRO QUECHUA	- (GUARANI is also NumN)	-	-
Penutian (3)	JACALTEC (Note that ZOQUE is also NumN)	?	?	?
other Amerindian (8)	CARIB LUISENO SHUSWAP	YUROK	?	?
Central NG (6)	-	- (MARIND and FORE also permit NNum)	-	DAGA KATE SENTANI WASKIA
Other NG (4)	-	SULKA=	-	SULKA= NASIOI SIROI

Not listed for lack of data:

GUARANI (Andean-Equatorial)
MAIDU, ZOQUE (Penutian)
GUAYMI, HIDATSA, POMO(E), ZAPOTEC (other Amerindian)
FORE, MARIND (Central NG)
OLO (other NG)

Here we find that only three major strata, Australian, Afro-asiatic and Sino-Tibetan have variation in their NumN/NNum ordering, apart from those languages, Malay, Ijo, Japanese, FORE and SUKA, which have both orders. Similarly only ~~three~~ four major strata, Austronesian, Niger-Kordofanian, Nilo-Saharan and Sino-Tibetan, have variation in the position of 'all', again not including families with languages which vary the position, ie. Kinyarwanda, Zulu, Berber, and Japanese. Only in the case of Sino-Tibetan do we find that such variation in the position of 'all' and of numerals is correlated. This situation may be summarised as follows:

$$\underset{(4)}{\text{allN/Nall}} \overset{1}{\longleftrightarrow} \underset{(3)}{\text{NumN/NNum}}.$$

This situation is a very interesting one when looking at post-positional languages: here a <u>low</u> correlation between allN/Nall and NumN/NNum within strata contrasts with a <u>high</u> correlation overall, ie. between strata. This suggests that postpositional languages only rarely change their ordering of 'all' and numerals, but when they do, they change either together or in rapid succession. If this is the case we might well expect something of this nature to be happening at present among the Sino-Tibetan languages, at least the post-positional languages in this group.

We may now compare the above pattern to that between allN/Nall and Pr/Po:

$$\text{Pr/Po} \longleftrightarrow \text{allN/Nall}$$

since the position of 'all' appears to be correlated with both numeral position and Pr/Po. We need only look at the strata for which there is variation in allN/Nall, and details for these appear in Table 9.7. (see page 345)

Here we also see the four strata that vary for position of 'all', though another ~~three~~ two strata had one language each varying the position of 'all' (~~Mandarin in Sino-Tibetan, Masai in Nilo-Saharan~~ Berber in Afro-Asiatic and Japanese in Ural-Altaic). Australian has been omitted on account of insufficient data. The data would appear to be problematical on account of the large number of languages which alternate their order for allN/Nall. But what is interesting is that one of the largest strata, Niger-Kordofanian, seems to provide a model to which others

approximate: although three of the languages, Kinyarwanda, Zulu and Ewe, have alternating or infixal ordering, the other six quite clearly correlate with Pr/Po: all the Pr languages except Sango place 'all' before the noun, while the two Po languages Ijo and Mandingo place it after.

This pattern is maintained with Sino-Tibetan and Nilo-Saharan, which tilt towards Po, and of the postpositional languages the larger number place 'all' after the noun. The Pr languages in these groups vary the position of 'all', as do Kinyarwanda and Zulu in the Niger-Kordofanian group. Likewise, if we look at Austronesian, which is exclusively prepositional, the distribution of orderings for 'all' corresponds broadly to that of the prepositional languages in Niger-Kordofanian.

Table 9.7 Partial Stratal Analysis

	allN		Nall	
	Pr	Po	Pr	Po
Austronesian (10)	Chamorro Malayan Palauan Tagalog Tsou Yapese	-	Cemuhi Fijian Maori Niue	-
Niger-Kordofanian (9)	Kinyarwanda= Zulu= Fulani Tiv Yoruba	-	Kinyarwanda= Zulu= Sango	Ijo Mandingo
Afroasiatic (6)	Berber= Cairene Hausa Iraqw Margi	-	Berber=	-
Nilo-Saharan (4)	Masai=	Kanuri	M̂asai=	Lugbara Kanuri
Sino-Tibetan (4)	Mandarin=	Newari	Mandarin=	Bawm Burmese
Ural-Altaic (3)	-	Japanese= Finnish Turkish	-	Japanese=

Not listed for infixing: Ewe (Niger-Kordofanian)

The above stratal analyses suggest that the position of 'all', allN/Nall, is correlated with both Pr/Po and NumN/NNum, though in different ways: when variation in the position of 'all' occurs within a stratum, this variation also correlates with Pr/Po within that stratum, but the correlation with NumN/NNum is basically discernible only across strata, and here only among postpositional languages.

Such a pattern suggests that we have a prediction chain of the form

$$\text{Pr/Po} \longleftrightarrow \text{all/Nall} \longleftrightarrow \text{NumN/NNum}$$

but this must be rejected. If we take the relevant data from Tables 9.4-5 we get the following figures:

Table 9.8

	Prepositional		Postpositional	
	allN	Nall	allN	Nall
NumN	12	5	9	1
NNum	6	2	1	12

As it stands the above data rejects the chain with Pr/Po in the middle, because of the correlation to be found under the Po columns. But if we re-arrange the table to put allN/Nall in the middle we get:

Table 9.9

	allN		Nall	
	Pr	Po	Pr	Po
NumN	12	9	5	1
NNum	6	1	2	12

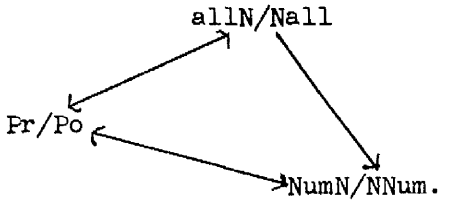
We may likewise draw up a third table:

Table 9.10

	NumN		NNum	
	Pr	Po	Pr	Po
allN	12	9	6	1
Nall	5	1	2	12

Whether we have NumN/NNum or allN/Nall in the middle, this must be rejected on account of the correlation in the right-hand column.

We can see then that no patterns of conditional independence may be set up for the above three orderings, and we have therefore a closed chain of the form



A similar analysis would show that Pr/Po also forms a closed chain with the remaining pairs:

NumN/NNum ⟷ smN/Nsm

and smN/Nam ⟷ NA/An

It is important to recall from earlier in the discussion that though the position of 'all' appears to correlate more closely with Pr/Po than with NumN/NNum, this does not apply to other elements in the hierarchy. Thus smN/Nsm correlates very closely with NA/AN, as may readily be seen from Tables 9.2-3.

The above discussion has served to show that it is not appropriate to place items on the adjectival hierarchy on a prediction chain. In particular it is extremely difficult to determine the precise role of NA/AN in the pattern of dependence and conditional independence on account of the few languages with AN/Nall, AN/NNum and AN/Nsm orderings. What we may observe though is that any attempt to form a chain consisting of Pr/Po, and two adjacent items on the adjectival hierarchy leads to a closed chain, with varying degrees of correlation within strata. Thus the position of 'all' and of numerals showed strong correlation among postpositional languages, but since each were slow to change, little correlation within strata occurred. Among exclusively Po strata however, when change did occur, as with the Sino-Tibetan languages, a clear correlation within the stratum also showed up.

This general phenomenon observed appeared to be due to three factors:

(i) the separate correlation of each item, allN/Nall, NumN/NNum, etc. with Pr/Po,

(ii) the strong correlation observed for such items among postpositional languages, and

(iii) the strong correlation of smN/Nsm with NA/AN, due either to (ii) above, or to the close syntactic and semantic correspondence of 'small' with other adjectival concepts, especially the benchmark adjectives 'big' and 'good' used in this present study.

The question now arises of how far a semantic explanation can be found for the above phenomena. But to answer this I must first

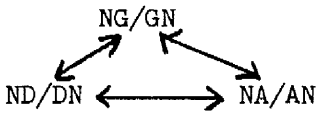
devote a section to the statistics of demonstrative position, and it will be clear from this discussion why among other things, I did not place demonstratives on the hierarchy.

9.2 The importance of deixis

In this section I wish to sort out the syntactic behaviour of demonstratives, and as it will turn out, pinpoint the source of certain syntactic patterns of the more central structures.

In chapter 1 I observed that demonstratives created a semantic region, and where possible, such a region was demarcated in the NP by the 'head' noun on the one hand, and the demonstrative on the other. Such an approach suggested that if adjectival and genitival modifiers were on the same side of the noun, the demonstrative would also appear on that side, perhaps between adjective and genitive NP.

Such a proposal would imply that the position of the demonstrative will be correlated with both genitival position, NG/GN, and adjectival position, NA/AN. This would correspond to the hypothesis that we have either a closed chain of the form



to be added somewhere to the adjectival hierarchy discussed above, or alternatively a straightforward prediction chain:

NG/GN ⟷ ND/DN ⟷ NA/AN

where it is only ND/DN that still depends upon the other two. In fact we can easily demonstrate that the latter is appropriate, with details in Table 9.10, and summary totals in Table 9.11 below. (See next page)

Table 9.10a

(Noun-demonstrative languages (ND))

	Noun-adjective (NA)	Adjective-noun (AN)
NG	Berber Cairene Cemuhi Fijian Fulani Iraqw (SOVPr) JACALTEC Kanuri (SOVPo) Malay Margi Niue Thai Tiv Vietnamese Welsh Yapese Yoruba ZAPOTEC (18)	-
NG ~ GN	OLO	-
GN	Bawm (also DND) Ewe Lugbara Wati GUAYMI HIDATSA DAGA KATE SIROI SULKA (SVOPr) WASKIA (11)	-

Table 9.10b

Demonstrative-noun languages (DN)

	Noun-adjective (NA)	Adjective-noun (AN)
NG	Kinyarwanda Mâsai (2)	Chamorro Greek Maung CAYUVAVA SHUSWAP (5)
GN	Abkhaz Burmese Mandingo Nubian GUARANI POMO, E. SENTANI (7)	Bardi Burushaski Finnish Hindi Ijo Japanese Kannada Mandarin (SVOPr) Nama Newari Tiwi (SVOPr) Turkish CARIB MAIDU QUECHUA YUROK (SVOPr) ZOQUE FORE MARIND (19)

Omitted from the above lists are Djingili (NG/NA) and LUISENO (GN/NA ~ AN) for which no information on position of demonstrative is available.

Maori (NG/NA), Zulu (NG/NA), Hausa (NG/AN) and NASIOI (GN/NA) are also omitted for ND ~ DN.

Also omitted are six other NA ~ AN languages, Sango (NG/ND), Palauan (NG/DN), Tsou (NG/DN), Ngandi (NG ~ GN/DN), PIRO (GN/DN) and Tagalog (NG/ND ~ DN).

Table 9.11

	Noun-demonstrative (ND)		Demonstrative-noun (DN)	
	NA	AN	NA	AN
NG	18	0	2(1.9)	5 (5.1)
GN	11	0	7(7.1)	19(18.9)

Here the bracketed figures under DN are the expected figures for when there is no direct correlation between NG/GN and NA/AN.

We can see clearly that any correlation in the sample between NA/AN and NG/GN can be explained by their respective correlations with ND/DN. That other orderings of the chain must be rejected may be seen from the following two tables.

Table 9.12

	Noun-adjective (NA)		Adjective-noun (AN)	
	ND	DN	ND	DN
NG	18(15.3)	2(4.7)	0	5
GN	11(13.7)	7(4.3)	0	19

In this table, under NA, there is a clear correlation between ND/DN and NG/GN which cannot be explained by respective correlations with NA/AN: under NA, most NG languages are ND, and most GN languages are DN.

Placing NG/GN in the middle of the chain we get:

Table 9.13

	Noun-genitive (NG)		Genitive-noun (GN)	
	ND	DN	ND	DN
NA	18	2	11	7
AN	0	5	0	19

Here we find a correlation between ND/DN and NA/AN which cannot be explained purely by their respective correlations with NG/GN. This is apparent among both NG and GN languages.

Now this chain is of especial interest, in that we have put together

three constructions, for each of which a different semantic account has been proposed: operator-operand for NG/GN, operator-upon-operator plus semantic region for ND/DN and finally adjunction for adjectives (NA/AN). A careful examination of this chain is therefore in order.

The most noticeable thing about the statistics from the above chain occurs when we look at the estimated percentages or probabilities: of the 33 DN languages, nine are NA (21.4%) and 24 are AN (66.7%). These are very similar distributions to those between ND/DN and NG/GN. Among the 31 ND languages however, all are NA, except for Sango, which has both NA and AN ordering.

The significant thing about such results is that it has been known for some time that NG languages are predominantly NA, whereas GN languages are more evenly divided, representing another asymmetric pattern. What neither Hawkins nor Greenberg attempted to do was look at demonstratives as thoroughly as they looked at NG/GN or NA/AN, though they were of course aware of the rarity of ND/AN languages. But what they could not do was tie the two facts together. This may be done by means of the prediction chain, showing that the prevalence of NA among NG languages is not due to the genitival ordering, but to the prevalence of ND languages among NG languages.

In looking for an explanation to the above patterns, with apparently three types of semantic structures involved, we must note that no AN/ND languages were to be found in the sample, and as far as I am aware, the only example of such a language outside the sample is that of Efik (NG/ND/AN), as shown in the following example:

àkámba	ébòt	òwò	òrò
	(ébót)	(ówó)	(órò)
big	goat	person	that

'That person's big goat'

(The items in brackets represent basic tonal contours.)

If we are to consider the possibility of an exceptionless universal being involved, we must clearly take Efik into account, and consider how constructions such as that above are exceptionless. We may note that adjective and noun and noun and genitive are juxtaposed in a very similar way, with tonal changes being triggered in the second item. In general the nature of the tone changes is determined by the tonal

contour of the first item, with slight differences in rules between nominal and adjectival first items. We may ask therefore whether the AN construction is in fact a noun plus genitive construction as suggested for Hausa, an AN language which permits ND as well as DN ordering.

The above question cannot really be resolved when we have only one example of such a language: we come back to the problems outlined in the first section of chapter 3, that in order to consider whether a universal is a candidate for treatment as an exceptionless one, we must firstly be clear as to what the relevant characteristics are, and secondly ensure that we are dealing with something that is statistically significant. In this regard we may note that Cemuhi, Welsh, Yapese and WASKIA permit a minority of items to precede which might be considered to be adjectival, despite their having ND ordering. We must therefore be clear as to precisely what we mean by the NA/AN typology.

Having focussed upon ND/DN as a crucial factor in the above asymmetric statistical patterns, we may ask how the phenomenon may be explained. We may firstly look at the NA/AN pattern: if adjectives are essentially adjuncts, why is it, according to my initial hypothesis, that they show any correlation with anything, especially as I suggest in chapter 1 that the motivating structure may be commutative?

The answer appears to be that the correlation arises not out of the nature of adjectives or adjuncts, but rather the nature of demonstratives, which need to mark out a semantic region, possibly in conjunction with other elements. It would appear however that the asymmetry can be captured by making different statements about the demarcation of the start of the region, and the demarcation of the end of the region. My proposal then is that the end boundary can be more easily inferred from the context than the starting boundary, when the context has yet to be expressed. Thus the following principle may be put forward:

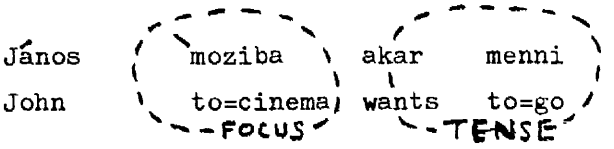
> Second-order operators may precede, follow or frame their semantic regions, but they ~~must precede~~ may not follow if the region begins with an adjunct.

It is important however to get this proposal into perspective: we have picked upon one example of a second-order operator, shown how

it differs semantically from first-order operators, and then shown how this particular example of a second-order operator appears to observe a different statistical pattern to other items, such as Pr/Po, NG/GN and VO/OV, which appear to be candidates for analysis as first-order operators. But for the proposal to hold, and avoid over-generalisation, we must ultimately consider other second-order operator constructions to see whether these follow a similar pattern.

Can we find this phenomenon at work then in other semantic regions - or at least hypothesised semantic regions?

One possibility is that of focus and tensed regions: an example of these, from Hungarian, would be:



'John wants to go to the cinema'

(eg. not to a play)

akin to the other Hungarian examples in chapter 1.

In Hungarian the focussed item comes either straight before the verb or is separated from it by negation. When such a focussed expression follows a topic (here 'János'), the first word, as observed in chapter 1, takes the sentence stress, accompanied by a high-fall accent (Kiss (1981)).

In English on the other hand, the focussed expression may follow the verb. In this case there is no need for the verb to come straight before the focussed expression in mirror image to the Hungarian pattern.

The treatment of tense as a second-order operator would at first sight appear to run counter to the approach of the tense logicians (see Allwood et al (1977), A.N. Prior (1967)), where tense is a first-order operator upon propositions. It may of course happen that natural language behaves entirely differently from formal logic, so that differences predict areas in which natural language will lead to false deductions in the area of tense. This would not be entirely surprising. However we must bear in mind that the logical form

Hp (Allwood et al convention)

has the meaning 'the proposition p has been true'. The question now becomes one of whether we can treat past tense (and indeed future tense) as a second-order operator with 'is true' as its region. If we can there would appear to be no incompatibility. But it is still necessary to proceed with caution: I am not making the claim that tense acts as an operator upon predicates such as 'is true'. Rather I have suggested that it operates upon a tensed region which interacts with the focussed item: this may well be the predicate, but it might equally well be only part of it.

I shall now turn to the example of <u>indefinite</u> determiners, and in my next section I shall argue both from semantic and from statistical evidence that they may be analysed either as first- or second-order operators.

9.3 Indeterminacy of semantic structure

In my last section I examined the behaviour of demonstrative position within the overall statistical pattern, and suggested that it was typical in its ordering of items creating semantic regions, ie. second-order operators. An important manifestation of this was its bar on AN/ND languages.

Now the difficulty about such an approach is that a similar bar appears, though not quite as strongly, to occur with the position of 'all' and that of numerals. Thus while we find that there were plenty of Pr/NA languages with allN or NumN ordering, there were comparatively fewer instances of the mirror-image situation, namely that in which Po/AN languages displayed Nall or NNum ordering, at least as their sole ordering. What then is the difference?

If we look at the data we see four basic differences:

(i) Although ND, NA, Nsm, NNum and Nall are all relatively consistent, as are DN, AN, smN, NumN and allN, ND differed from Nsm, NNum and Nall in being relatively consistent with Pr, rather than Po. A similar statement could be made about DN. To illustrate this, we may observe that among prepositional languages there was a tendency for numerals and 'all' to precede their noun, whilst demonstratives tended to follow.

(ii) Among Po/NA languages, almost all had Nall and NNum ordering. Only some 30% had ND ordering.

(iii) The bar on AN/ND ordering appears to be almost absolute: the only counter-example I am aware of is Efik. By contrast it is not difficult to find AN/Nall languages: Burushaski and Ijo (Po) for example. Nor AN/NNum languages: thus Hausa (Pr) and MARIND (Po). Furthermore a number of Po/AN languages had Nall or NNum as alternative orderings.

(iv) Different prediction chain patterns were observed. Thus ND/DN entirely explained the correlation between NA/AN and Pr/Po. This did not occur with other items, such as allN/Nall, NumN/NNum or smN/Nsm, where the pattern of a Möbius band occurred.

The account I would propose for this situation, in which items on the adjectival hierarchy, at least 'all' and the numerals, partly correlate

with Pr/Po and partly follow ND/DN in avoiding postposing in AN languages, is one in which indefinite determiners vary in their interpretation as first-order or second-order operators.

We can readily see this in the following example:

four grandsons of the mayor,

four sons of a son/daughter of the mayor.

Here the numeral 'four' must govern a region, since the four sons concerned cannot be any four grandsons, but those of a particular son or daughter of the mayor. But this approach does not apply in the expression

four of the sons of a son/daughter of the mayor.

Although the example has the same meaning as the previous one, beginning 'four sons...', it is not the numeral which creates the region, but the definite article, which has now been introduced. We may now treat the expression 'four (of)' as a first-order operator.

We can similarly contrast 'all sons' with 'all the sons' and 'all of the sons', and as a result we have an explanation for why items such numerals and certain indefinite determiners are subject to the constraints of both first- and second-order operators.

9.4 Summary

In this work I have attempted to separate out a number of types of statistical patterns and to identify them with a number of semantic patterns, and I would like in this section to list the word-order principles which have emerged and their associated problems.

Essentially there are three principles, one each for first-order operators, second-order operators and adjuncts, though the last two involve the interaction of more than one type of structure:

(1) First-order operators will tend to consistently precede their operands within a language, or consistently follow.

(2) Second-order operators may precede, follow or frame their semantic region, but they ~~must precede~~ may not follow if the region starts with an adjunct.

(3) Adjuncts will tend to follow their adjoinee expression if their internal structure places a head item at the beginning, and to precede if their internal structure places a head item at the end.

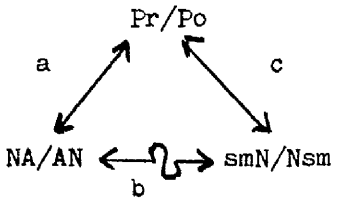
Looking firstly at principle (1), we may note that this is simply a carry-over from the Bartsch/Vennemann Principle of Natural Serialisation. The critical change in this work is that if we take intensional adjectives as a yardstick for what is an operator as opposed to what is an operand, we must swop round virtually all the Bartsch/Vennemann judgments as to what constitutes an operator and what constitutes an operand. In particular non-subject NPs must be treated as operands, unless of course there are grounds for treating them as adjuncts to the verbal complex (see chapter 8).

The manifestation of this re-adapted principle of natural serialisation is most apparent in the chain

$$VO/OV \longleftrightarrow Pr/Po \longleftrightarrow NG/GN.$$

Such a chain suggests that operator-operand structures correlate most strongly with Pr/Po, and we also observed the correlation of allN/Nall, NumN/NNum and smN/Nsm with Pr/Po. However the evidence from chapter 4 suggested that when case-marking occurred, this latter took over as the central factor to which operator-operand structures correlated.

An extremely important statistical pattern in this regard is the closed chain of the form:



This pattern is of particular importance, with its Möbius twist at b, in that it cannot be explained solely by a single explanatory factor, such as the Natural Serialisation Principle. Bartsch and Vennemann only considered the link a in their supposedly simplified model in which all adjectives were treated alike as Operators. Consequently prepositional object NPs were also treated as Operators. But in fact the only link which can be so treated is c, leading to my re-adapted principle under which it is prepositions rather than prepositional object NPs that are treated as operators.

Consequently we must find some other explanation for links a and b. Since b requires little investigation on account of the similar syntactic treatment of 'small' and other adjectives, I shall concentrate upon a.

To account for a, principle (2) becomes relevant. Here we must expand the link to give

$$\text{Pr/Po} \longleftrightarrow \text{NG/GN} \longleftrightarrow \text{ND/DN} \longleftrightarrow \text{NA/AN}$$

It must be stressed here that the ordering of Pr/Po and NG/GN comes from the Hawkins sample: data from my own sample was insufficient to confirm the ordering. Now the above principle becomes relevant because I have suggested that the behaviour of demonstratives could be explained by their status as second-order operators. But what we have in reality is two separate phenomena, one statistical and one semantic, which have been married here for the sake of discussion.

The statistical phenomenon is that of the chain

$$\text{NG/GN} \longleftrightarrow \text{ND/DN} \longleftrightarrow \text{NA/AN},$$

where the ban on AN/ND languages has parallels in the ordering of 'all' and the numerals. This matter requires further investigation in that the link between genitival ordering (NG/GN) and demonstrative

ordering (ND/DN) needs to be more fully probed.

The semantic phenomenon is of course that of the second-order operator, and further investigation is certainly required on this count. Until that is done though, we have a problem in that demonstrative ordering was one example of a particular statistical pattern, and the one example of a particular semantic structure, apart from certain parallels in the position of 'all' and the numerals. For this reason, we cannot strictly say that the semantic pattern explains the statistical pattern, so much as coincides with it.

We now arrive at principle (3), relating the internal and external syntax of adjuncts. This was illustrated only with respect to case-marked subjects and objects with respect to their verb. However this is a very plausible principle on other counts: in particular we find that OV languages tend to place their relative clauses before the noun. Principle (3) predicts this, since the verb, as a head element in the subordinate clause, comes at the end. In addition pre/post-positional phrases (PP) will tend to follow the verb in Pr languages and precede in Po languages. It is important to note that when principle (3) is introduced, fusion of case-marking as discussed in chapter 4 represents a violation of that principle, as long as we treat case-marked NPs as adjuncts to the verb.

It is important to note the difference between Principle (3) and the Structural Principle of Language proposed in Lehmann (1973), whereby 'modifiers are placed on the opposite side of a basic syntactic element from its primary concomitant'. Here the basic syntactic element parallels the head of an adjunct expression, eg. the preposition in the PP, and the primary concomitant likewise parallels the adjoinee expression. However the concept of adjoinee is much more specific. In this way we have the interaction of three items. The difference is of course that my Principle (3) strictly applies to adjunct expressions only.

It will be seen then that a system distinguishing first-order operators, second-orders and adjuncts has strong support from the statistics considered in the present study. It provides a more attractive account than either the Hawkins or Bartsch/Vennemann approach, and is well-suited for the basis of a wider study.

Footnotes

to

Part II

Footnotes for Chapter 6

1. Siegel in fact uses the terms 'intersectional' and 'non-intersectional' (see later in this section for further mention). These correspond respectively to 'extensional' and 'intensional' and I have stuck to the latter for the sake of consistent usage. It must however be borne in mind that the treatment of non-intersectional adjectives as intensional is a feature of intensional logic. It is a matter of demonstration whether such treatment has any linguistic reality.

2. Such a class of adjectives might include those identified by Grevisse (1955) as following the noun in French: 'Les adjectifs dérivés d'un nom propre et ceux qui marquent une catégorie religieuse, sociale, administrative, technique, historique, géographique, etc.'

3. We may note however that (a) Lexical and (d) non-inherent Semantic, have been tied up with the intensional/extensional dichotomy: Siegel (1980) considers this with respect to Italian attributive adjectives. She considers post-nominal restrictive adjectives to be extensional, in effect restricting the extensions of the nouns they modify:

   inglesi flemmatici, Englishmen who are phlegmatic (RESTRICTIVE)

   flemmatici inglesi, Englishmen, who are phlegmatic (NON-RESTRICTIVE)

   On the other hand pre-nominal adjectives may in some instances take on a 'senso traslato', a metaphorical or 'transferred' meaning, which does not occur predicatively. These are considered to be intensional meanings, so that 'buono' means 'kind' when it follows (extensional), but 'good' when preceding: buon professore, good (as a) professor, rather than a 'kind professor'. Such adjectives may in fact be restrictive when occurring pre-nominally with senso traslato.

4. See Whitney (1956), Notes 14(e). For further discussion of Finnish in this study, see § 6.2.3.

5. For this example see Tucker et al (1955) para 259, where tonal case-agreement is illustrated. The rules illustrated indicate why the tone pattern of the noun varies in accordance with whether it precedes or follows the adjectives.
   See also paragraphs 22, 277-8 for so-called indefinites, including hoo, any.

6. See Briggs (1961) chapter 4 on 'Statives': morpheme 85 is added when the stative is used predicatively, to denote a 'state of being, a descriptive quality', and among the statives listed we find rob, big, listed as taking allomorph 85.9 gu-, and wên, good, listed as taking allomorph 85.13, the zero prefix. Note also that the na- of nayɨhL, is cool, is the allomorph of 85.1.
   Note that -gai, male, and -rô, big, are discussed in chapter 5 on 'Nouns': here they are listed as bound morphemes, numbered 414 and 426 respectively. In this same chapter the morpheme 541, s-, another, is also discussed.

7. Under A, B and C I have only listed the four languages Cemuhi, Mâsai, SULKA and ZAPOTEC. Other languages do of course have preceding A, B or C items, but they are discussed below. Under A we may also add ZAPOTEC, which expresses 'other' by the prefix s-,

7.(cont'd)

but this does not appear to be part of the adjectival class I have discussed for ZAPOTEC above.

8. The status of 'usual' as D(i) is problematic, despite the fact that it appears to have temporal content: if I am the usual night-porter, this does not necessarily mean that I am usually the night porter, as opposed to say sleeping, but that the night-porter is usually me, as opposed to anyone else.

An alternative approach might put 'usual' together with ordinals (C): neither the first nor the last, but the usual, or perhaps the average or median. But a difference of C or D(i) is not going to affect the overall argumentation of these chapters.

8a. Note the sources for the examples in the above and following pages, for JACALTEC, Tiv and Welsh.

For JACALTEC, see Craig(1977), chapter 1 on 'Typological Characteristics', but for the example translated as 'the river bank', see Chapter 2, example (11).

For Tiv, see Abraham (1940), chapters II (esp. §14), III (esp. §30), IV, XVIII and XXXIII.

For Welsh, I would like to thank John Wells for his help in compiling this list of adjectives, drawn partly from James, D.L. (1972).

9. See Milner (1956). See paragraphs 73 and 80. Example taken from Exercise 12 II 6.

10. 'Auxiliary nouns' include cì, big, very, and (hkǎ)leì, (a) little, but little else that could be considered to be descriptive adjectives (see Okell (1969)):

A	taìn˜	every
B	toú, tei,	
	twei, myà	Plural
C	mà	main, feminine (this item pc, Okell)
	hsoùn˜	extreme
D	lau˜	approximately
E	(see above)	

Other 'auxiliary nouns' only follow numerals.

11. See Westermann (1930) para 76 item 5.

12. These two GUARANI examples are to be found in Gregores and Suárez (1967) in sections 15.434 and 13.21232.

13. See Spears (1965), chapter V, 5.4 on 'Transformational Rules': the example given illustrates 160 (p.101). The discrepancies in the tone contours are due to the fact that I have attempted to apply rules 169, 178 and 179 (see section 5.5 on 'Phonetic Terminal Output') to yield a surface structure. It may well be that other rules apply as well, once a sentential context has been found for this postpositional phrase.

14. I am grateful to Mr Haroun Rashid for his discussion of these examples, and in particular for supplying me with the examples illustrating the rendering in Hausa of old, former, fast and adjectives of material and nationality.
The other Hausa examples come from chapters in Kraft and Kirk-Greene (1973), illustrating specifiers, indefinite nominals and adjectival nominals respectively: the example 'some big aeroplane' is drawn from the exercises in chapter 23 and its English translation is to be found in the Key.

15. See Topping (1973) section 4.2 on '(Syntax of) Modifiers of the Noun'.

16. See Drabbe (1955), chapters II-III. In particular see paragraphs 16-17, 19, 23-25. (Para. 25 discusses an alternative construction for 'all'.)

17. The examples of the expression of all, every in Japanese come from two separate sources, Dunn and Yannada (1958) and Kuno (1972).
The example illustrating dono...mo appears in chapter 28 of Kuno (1972). The example given implies that in every pond there were kinds of fish, but not necessarily the same kinds, as could be implied if the dono...mo had come first.
The example illustrating mina comes from my principle source, Dunn and Yannada (1958), from chapter 17, where it is stated that certain adverbs, including mina, represent the most satisfactory way of translating expressions involving 'all', 'every', etc. when they are not particularly emphasised.
I am grateful to Sho Jiro Sekine for discussion of mina and the possibility of ambiguity in its use.

18. See Williamson (1965), sections 2.2 and 4.5 for examples.

19. Discussion of these two items occurs in Hoff (1968), in section 3.2.1.2. on Nominal Word-Groups, in a footnote. The first example appears in this same footnote, and indeed in Text 4.4. The second example occurs in Text 4.2 (p.297).

20. There is evidence in both Maung and Tiwi that expressions of totality are attached to the verb. Thus in Maung we have the indefinite pronoun girg~jirg which follows the verb of which it is subject or object.

Note the following Maung example, taken from Cap ell and Hinch (1970), Text 3 towards the end:

(Mararawg aniwuŋ da warjad)
lightning hit the rock

awuniwuŋ girg dja mararawg bada waramidjmidj
hit(them) all the lightning the sandflies

'(Lightning hit the rock and) the lightning
hit all the sandflies.'

Here the girg is separated from the rest of object, bada waramidjmidj, the sandflies.

In Osborne (1974) we find the verbal infix -tɔmɔnti-, glossed as 'all' in Tiwi:

jiŋwampa woŋatuŋa pu- tɔ- tɔmɔnti- kɔɻimu
others basket 3pl Pa all make

'The others were all making baskets.'

As far as I can tell, this infix is listed in neither the grammar nor the dictionary of Osborne's work. Note also that this example is not an SOV construction: instead 'basket' combines with the ancillary verb 'make' to form a periphrastic lexical item.

21. It is not clear how numerals are dealt with in MAIDU from my main source, Shipley (1964). However according to Boas (1911), numerals are a sub-class of adjectives, which in MAIDU precede the noun with the suffix -Im (also used for the nominative case).

Neither indicate how 'all' is treated.

FOOTNOTES FOR CHAPTER 7

1. See Ruffell Barlow (1951); Study XII for example

2. See chapters VII-VIII of Tung (1964). Chapter VII describes the 'core', which is essentially the verb complex, while chapter VIII describes 'Conjunctive Phrases' which appear to correspond to non-predicate NPs.

   The first example is to be found in VIII 6 (on 'Subordinative Conjunctive Phrases, Type 1') and the second in VII 3. The explanation of the use of no as in 'méoi no cou', a big man, occurs partly in VII 3 and partly in VIII 7 ('Type 2').

3. A note of caution is appropriate here: Greenberg classifies GUARANI as GN. Normally I have considered the placing of genitival NPs. However Gregores and Suárez (1967) only illustrate genitival nouns. Thus the grammatical section 15.1, outlining nominal phrases has:

   kwati.á pepó
   paper wing

   tah+í rait+
   ant(s) nest

   which looks strongly like English N+N constructions rather than possessive constructions. The example for 'bank of a stream' is taken from Text 4 (Sentence 1) and the 'strong wind' example from Text 2 (Sentence 10).

4. The Burmese example is taken from Okell (1969), chapter 3 section 29.
   The PIRO example is taken from section 232 of Matteson (1965).

5. See Williams (1923), § 87, for translation of and.

6. See Schachter and Otanes (1972), §4.12 on 'Intensification of Adjectives', pp. 234-5.

7. See Schachter and Otanes (1972), § 5.7 on 'Benefactive-focus Verbs' for example.

FOOTNOTE TO CHAPTER 9.

1. Note discussion of such implications in chapter 6 (§6.1).

Conclusion

In concluding my work I wish to list the areas in which I consider that an original contribution has been made. In doing this I shall consider three areas:

(i) the statistical findings,

(ii) the semantic findings,

(iii) the discussion linking the two, and in particular leading to the formulation of the three principles or constraints outlined in §9.4.

The statistical findings comprise five main points:

(i) the statistical non-significance of certain of Hawkins's proposed exceptionless or 'implicational' universals (see §3.2).

(ii) the alternative analysis of Hawkins's data, involving patterns of conditional independence, stated in the form of the prediction chain (§3.3) and yielding the High Odds and Evens Principles (§3.5).

(iii) the stratal analysis of the data in my own sample, revealing differences, perhaps diachronic, in certain of the word-order patterns (§2.3), and yielding the principle of the Diminished Trigger Effect (§3.5).

(iv) the greatest incidence of fusional languages among inconsistent languages in terms of the VO/OV-Pr/Po correlation (chapter 4).

(v) the conflation of the correlations of adjectival position with NG/GN and with ND/DN by means of the prediction chain (§9.2).

It is important to note that the concepts of conditional independence and of stratal analysis are not new. What is important is their application to problems for which Hawkins proposed 'implicational' (ie. exceptionless) universals, and for which Vennemann proposed what Hawkins terms a 'Trigger Chain'. In effect Hawkins' Universal Consistency Hypothesis is replaced by the combination of the High Odds and the Evens Principles, while his Universal Violation ('Trigger Chain') Hypothesis (UVH) is replaced by the Diminished Trigger Effect. Unlike the UCH and UVH however, the replacement hypotheses are not

incompatible for any given construction.

I shall now turn to the semantic findings, of which there are three:

(i) Analysis of the word-order pattern of individual adjectives, so that intensional adjectives are isolated, requires a re-adaptation of the Bartsch/Vennemann Natural Serialisation Principle, so that in general non-subject NPs become operands rather than operators (chapters 6-7).

(ii) The adjunct relationship is identified as a distinct semantic concept (chapter 7).

(iii) Separate forms of operator, first- and second-order, are distinguished in the operator-operand analysis of certain constructions (chapters 1 and 9).

The first finding is a very important one, and is also a welcome one, in that concepts such as 'brother of', have normally been treated as operators rather than operands, along with mathematical concepts such as 'square of'. We are now able to do this directly, since the examination of intensional adjectives, that is those which most need to be treated as operators, shows that when they differ from other adjectives in their ordering, they are more likely to precede in Pr languages and to follow in postpositional languages. As a result prepositions and postpositions may also be treated as operators in as far as word-order is to be explained by an operator-operand account.

The second finding is a continuation of the approach of Kamp (1975) and Siegel (1980) in which some adjectives which should strictly be treated as intensional are in fact treated as extensional with allowance for context. In this group Siegel includes 'measure' adjectives, which may include value, dimension and speed adjectives. My own approach however, isolates certain adjectives, generally intensional, as operator, and treats much of the rest as 'adjuncts'. Among such adjectives there is no need to treat the 'good' in 'good pick-pocket' as intensional: the adjunct refers to the most relevant concept embodied in the adjoinee expression.

The identification of the adjunct relationship as a semantic concept permits us to put forward the hypothesis that the operand in an operator-operand structure is the item which may be expanded to include adjuncts, though this posed problems, discussed in chapter 8,

if we wished to treat nominal subject NPs as operators.

The idea of distinguishing first- and second-order operators is not a new one. Logicians who treat adverbs as predicates of predicates are in fact treating such adverbs as second-order operators, in as far as predicates are operators upon their arguments (see Allwood, Andersen and Dahl (1977)). However I believe my semantic and statistical account of definite determiners and sometimes indefinite determiners as second-order operators does represent a new contribution.

In the above statistical and semantic findings I have not included the three principles outlined in §9.4. Principle (1) states the consistency of operator-operand orderings. But this is simply the Natural Serialisation Principle, re-adapted to a different concept of what constitutes operator and operand. Principle (2) would at first sight appear to be new: the difficulty is that the constraint is heavily dependent upon the statistical pattern of demonstratives and therefore requires closer investigation of the concept of second-order operator. Principle (3) applies strictly to the behaviour of adjunct expressions, and has an interesting parallel to Lehmann's Structural Principle (see §9.4). But even though there appears to be ample evidence for it, it has in this work strictly only been investigated in connection with case-marked subject and object: in this regard it appears to represent the rule whose violation motivates morphological fusion.

It will be seen then that this study has yielded important results in the area of statistical investigation, revising both the Universal Consistency Hypothesis and the Trigger Chain account. It further showed that the Natural Serialisation Principle needed to be re-adapted in the light of a study of adjectives in which differences in types of adjectives were investigated, rather than generalisations sought.

The resulting discussion suggested that important differences could be found in the types of operator, first- and second-order, with different patterns of word-order. In regard to second-order operators, the semantic properties of determiners, and of tense and focus, would appear to be an interesting area for further investigation.

APPENDIX I. LANGUAGES AND REFERENCES

In this Appendix I list languages in alphabetical order, giving their geographical area and references used in this study. The first reference listed may be referred to for relevant information, including grammatical points and examples, unless otherwise indicated in the text or footnotes. (Except of course English, where the examples and judgments may in fact be my own.)

All references are from Bibliography B unless the date is followed by '(A)'.

Language	Geographical Area	Reference(s)
Abkhaz	Asia, Caucasus	Hewitt, B.G. (1979)
Akkadian	(Babylon)	Von Soden, W. (1952)
Amharic	Africa, Ethiopia	Armbruster, C.H. (1908)
Arabic (Colloquial Egyptian)	see Cairene	
Arapesh	New Guinea	Fortune, R.H. (1942)
Bardi	Australia, Western	Metcalfe, C.D. (1975)
Bawm	Asia (South) Bangladesh	Reichle, V. (1981)
Berber	Africa, North	Aspinion, Robert (1953)
Burmese	Asia, S.E.	Okell, J. (1969)
Burushaski	Asia (South), Kashmir	Lorimer, D.L. (1938)
Cairene	Africa (North), Egypt	Mitchell, T.F. (1956)
Carib	America (South), Surinam	Hoff, B.J. (1968)
Cayuvava	America (South), Bolivia	Key, H. (1967)
Cemuhi	Pacific, New Caledonia	Rivierre, J-C. (1980)
Chamorro	Pacific, Mariana Is. (incl. Guam)	Topping, D. (1973)
Chinese (Mandarin)	see Mandarin	
Daga	New Guinea	Murane, E. (1974)
Djingili	Australia, NT	Chadwick, N. (1975)
Efik	Africa (West), Nigeria	Welmers, W.E. (1968)
English (Modern)	Europe, Western, etc.	Quirk et al (1972) Creider, C.A. (1979)(A) Montague, R. (1970a, b) (A)

Appendix I (continued)

Language	Geographical Area	Reference(s)
English (Old)	(Europe, Western)	Canale, W.M. (1978)
Eskimo	Arctic	Comrie, B. (1981b) (A)
Estonian	Europe, North (Baltic)	Comrie, B. (1981a) (A)
Ewe	Africa (West), Ghana	Westermann, D. (1930)
Fijian	Pacific	Milner, G.B. (1956)
Finnish	Europe, North (Baltic)	Whitney, A.H. (1956)
Fore	New Guinea	Scott, G. (1978)
French	Europe, West, etc.	Grevisse, M. (1955) Ferrar, H. (1967)
Fulani	Africa, West	Arnott, D.W. (1970)
German	Europe, West	Bartsch, R. (1976) (A)
Greek (Modern)	Europe, S.E.	Sofroniou, S.A. (1962)
Guaraní	America (South), Paraguay	Gregores and Suarez (1967)
Guaymí	America (C), Panama	Alphonse, E.S. (1956).
Hausa	Africa, West	Kraft and Kirk-Greene (1973)
Hidatsa	America (North), Dakota	Matthews, G.H. (1965)
Hindi	Asia, South	McGregor, R.S. (1977)
Hungarian	Europe, Central	Banhidi et al. (1965) Tompa, J. (1968) Kiss, K.E. (1981)
Ijo	Africa, West	Williamson, K. (1965)
Iraqw	Africa (East), Tanzania	Whiteley, W.H. (1958)
Italian	Europe, South	Siegel, M.E.A. (1980)(A)
Jacaltec	America (C), Guatemala	Day, C. (1973) Craig, C.G. (1977)
Japanese	Asia, Far East	Dunn and Yannada (1958) Kuno, S. (1972,1973,1978)
Kannada	Asia, South	McCormack, W. (1966)
Kanuri	Africa (C), Lake Chad	Lukas, J. (1937)
Kashmiri	Asia (South), Kashmir	Grierson, G.A. (1911)
Kate	New Guinea	Schneuker, C.L. (1962)
Khamtî	Asia (South-east), Burma	Needham, J.F. (1894)
Kikuyu	Africa (East), Kenya	Ruffel Barlow, A. (1951)
Kinyarwanda	Africa (Central) Ruanda	Kimenyi, A. (1978)

Appendix I (continued)

Language	Geographical Area	Reference(s)
!Kung	see !Xu	
Latin	(Mediterranean)	Ramsay, G.G. (1897)
Latvian	Europe, North (Baltic)	Lazdina, T.B. (1966)
Lugbara	Africa (NE), The Sudan	Crazzolara, J.P. (1960)
Luiseno	America (N), California	Kroeber and Grace (1960)
Maidu	America (N), California	Shipley, W.F. (1964) Boas, F. (1911)
Malay	Asia, South-east	Dodds, R.W. (1977)
Mandarin	Asia, Far East	Chao, Yuen-ren (1968) deFrancis, C. (1965) Li and Thompson (1973, a,b, 1975)
Mandingo	Africa, West	Spears, R.A. (1965)
Maori	Pacific, New Zealand	Williams, W.L. (1923) Biggs, B. (1973)
Margi	Africa, West	Hoffman, C. (1963)
Marind	New Guinea	Drabbe, P. (1955)
Mâsai	Africa (E), Kenya	Tucker and Tompo Ole Mpaayi (1955)
Maung	Australia, Western	Capell and Hinch (1970)
Nama	Africa (Southern), Namibia	Hagman, R.S. (1977)
Nasioi	New Guinea	Rausch, P.J. (1912)
Newari	Asia (South), Nepal	Sresthacharya, I. et al (1971)
Ngandi	Australia, NT	Heath, J. (1978)
Niue	Pacific	Seiter, W.J. (1980)
Nubian	Africa (NE), The Sudan)	Lepsius, R. (1880)
Olo	New Guinea	McGregor and McGregor (1982)
Palauan	Pacific, Palau	Josephs, L.S. (1975)
Pashto	Asia (C), Afghanistan	Penzl, H. (1965)
Persian	Asia	Mace, J. (1962)
Piro	America (S), Peru	Matteson, E. (1965)
Pomo, Eastern	America (N), California	McLendon, S. (1975, 1978)
Quechua (Ayacucho)	America (S), Peru	Parker, G.J. (1969)
Russian	Europe, North	Semeonoff, A. (1955)
Rutul	Asia/Europe, Caucasus	Dirr, A. (1911)
Sango	Central African Republic	Samarin, W.J. (1967)

Appendix I (continued)

Language	Geographical Area	Reference(s)
Sanskrit	(South Asia)	Coulson, N. (1976)
Sentani	New Guinea	Cowan, H.K.J. (1965)
Shuswap	America (N), British Columbia	Kuipers, A.H. (1974)
Siroi	New Guinea	Wells, M.A. (1979)
Songhai	Africa (W), Mali	Hacquard and Dupuis (1897)
Spanish	Europe, West, etc.	Ramsey, M.M. (1956)
Sulka	New Guinea	Müller, H. (1916)
Swahili	Africa, East	Ashton, E.O. (1947)
Tagalog	Pacific, Phillipines	Schachter and Otanes (1972)
Thai	Asia, South-eastern	Noss, R.B. (1964)
Tigrinya	Africa, North-east	Praetorius, F. (1871)
Tiv	Africa, West	Abraham, R.C. (1940)
Tiwi	Australia, NT	Osborne, C.R. (1974)
Tsou	Asia (Far East), Taiwan	Tung, T'ung-ho (1964)
Turkish	Middle East	Lewis, G.L. (1953) Ergunvanlı, E.E. (1984) Lewis, G.L. (1967) Underhill, R. (1976)
Vietnamese	Asia, South-east	Thompson, L. (1965)
Waskia	New Guinea	Ross, N. (1978)
Wati	Australia, WA/SA	Douglas, W.H. (1959)
Welsh	Europe, Western	Bowen and Rhys-Jones (1960)
!Xu	Africa (SW), Kalahari	Snyman, J.W. (1970)
Yapese	Pacific, Yap	Jensen, J.T. (1942)
Yoruba	Africa, West	Rowlands, E.C. (1969)
Yurok	America (N), California	Robins, R.H. (1958)
Zapotec (Mitla)	America (N), Mexico	Briggs, E. (1961)
Zoque	America (N/C), Mexico	Wonderley, W.M. (1952)
Zulu	South Africa	Rycroft and Ngcobo (1981)

APPENDIX II. LANGUAGES AND THEIR TYPOLOGY

In this Appendix I list the languages of the sample in this present study, their typological characteristics in terms of VSO/SVO/SOV, Pr/Po, NG/GN and NA/AN, and in the right-hand column, a geographical/genetic classification. (In fact, this is one used at one stage in Chapter 2, and the distinction is purely geographic, except for the identification of Austronesian).

Some comments are appropriate on the typological criteria used.

All the languages have been characterised for Pr/Po, VS/SV and VO/OV, the latter two implicit in the VSO/SVO/SOV classification (with some VOS and one possible OVS). In certain cases however a VSO/VOS distinction was by no means obvious. In the case of Tagalog and Tsou I found it inappropriate to make any judgment (see Chapter 4). I have however followed Greenberg and Hawkins in treating Maori as VSO: this it undoubtedly is in active clauses, but with transitive verbs the passive voice appears frequently to be the unmarked usage. Metcalfe (1975) treats Bardi as having a basic VSO order in his transformational account, with the preverbal position used for the 'topic'. In the texts provided, however, I found no single example of a clause with both nominal subject and object.

Although the basic analysis for PIRO in Matteson suggests an SOV order, a large number of texts are provided, and VS occurs frequently, often for a definite subject. My assignment of OVS is highly tentative, and as much as anything else, experimental. It could be argued that since nothing significant emerges from the statistics in treating it as either VS or SV, we are in no position to usefully refine our VS/SV criteria to deal with languages such as PIRO.

Since there are no languages such as Lithuanian or the Algonquian languages in the sample, there is little need to add to the discussion on Pr/Po in Chapter 3 with special reference to SVO languages. Similarly I discuss the question of classifying languages for NA/AN in Chapter 6.

However NG/GN did provide a serious problem, in particular with the languages Bardi, Maung, Ngandi, OLO and Tiwi. It seemed clear to me that Ngandi could not be assigned a dominant order for NG/GN, and I could not usefully add to Heath (1978)'s discussion on Ngandi. OLO appeared to use entirely different genitive constructions, depending on whether or not the NP concerned was a direct object - the main construction exemplified - and it seemed inappropriate to assign any dominant ordering. In the case of Bardi, Maung and Tiwi there appeared from examples to be a marginal preference for one ordering, NG for Maung and GN for Bardi and Tiwi, and I made these choices to distinguish them from Ngandi and OLO. Although Greenberg and Hawkins treat Tagalog as NG ↭ GN, my reading of Schachter and Otanes (1972) strongly suggested that NG was the dominant ordering, with the 'N ng G' construction

the preferred one.

Since my assignment of NG/GN was as even as it could be, with the odd number of three problematic languages: Bardi, Maung and Tiwi, it is unlikely that these assignments will seriously affect the present statistical analysis. Nevertheless it will be recalled that the choice of NG/GN as intermediary on the prediction chain between Pr/Po and VS/SV was highly precarious, so that in future research, a clearer definition of the NG/GN typing might well be valuable.

It will be noted that Greenberg and Hawkins assign AN typing to Abkhaz: a reading of Hewitt (1979) shows that this is clearly incorrect.

Language	Basic typology				Geographical /Genetic
Abkhaz	SOV	Po	GN	NA	Asia
Bardi	VSO?	Po	GN?	AN	Australia
Bawm	SOV	Po	GN	NA	Asia
Berber	VSO	Pr	NG	NA	Africa
Burmese	SOV	Po	GN	NA	Asia
Burushaski	SOV	Po	GN	AN	Asia
Cairene	SVO	Pr	NG	NA	Africa
CARIB	SOV	Po	GN	AN	America (S)
CAYUVAVA	VOS	Pr	NG	AN	America (S)
Cemuhi	VOS	Pr	NG	NA	Austronesian
Chamorro	SVO	Pr	NG	AN	Austronesian
DAGA	SOV	Po	GN	NA	N. Guinea
Djingili	SOV	Po	NG	NA	Australia
Ewe	SVO	Po	GN	NA	Africa
Fijian	VOS	Pr	NG	NA	Austronesian
Finnish	SVO	Po	GN	AN	European
FORE	SOV	Po	GN	AN	N. Guinea
Fulani	SVO	Pr	NG	NA	Africa
Greek (Modern)	SVO	Pr	NG	AN	European
GUARANI	SVO	Po	GN	NA	America (S)
GUAYMI	SOV	Po	GN	NA	America (C)
Hausa	SVO	Pr	NG	AN	Africa
HIDATSA	SOV	Po	GN	NA	America (N)
Hindi	SOV	Po	GN	AN	Asia
Ijo	SOV	Po	GN	AN	Africa
Iraqw	SOV	Pr	NG	NA	Africa
JACALTEC	VSO	Pr	NG	NA	America (C)
Japanese	SOV	Po	GN	AN	Asia
Kannada	SOV	Po	GN	AN	Asia
Kanuri	SOV	Po	NG	NA	Africa
KATE	SOV	Po	GN	NA	N. Guinea
Kinyarwanda	SVO	Pr	NG	NA	Africa
Lugbara	SVO	Po	GN	NA	Africa
LUISENO	SVO	Po	GN		America (N)
MAIDU	SOV	Po	GN	AN	America (N)
Malay	SVO	Pr	NG	NA	Austronesian

Language	Basic typology				Geographical /Genetic
Mandarin	SVO	Pr	GN	AN	Asia
Mandingo	SOV	Po	GN	NA	Africa
Maori	VSO?	Pr	NG	NA	Austronesian
Margi	SVO	Pr	NG	NA	Africa
MARIND	SOV	Po	GN	AN	N. Guinea
Mâsai	VSO	Pr	NG	NA	Africa
Maung	SVO	Pr	NG	AN	Australia
Nama	SOV	Po	GN	AN	Africa
NASIOI	SOV	Po	GN	NA	N. Guinea
Newari	SOV	Po	GN	AN	Asia
Ngandi	SOV	Po			Australia
Niue	VSO	Pr	NG	NA	Austronesian
Nubian	SOV	Po	GN	NA	Africa
OLO	SVO	Pr		NA	N. Guinea
Palauan	SVO	Pr	NG		Austronesian
PIRO	OVS?	Po	GN		America (S)
POMO, E	SOV	Po	GN	NA	America (N)
QUECHUA	SOV	Po	GN	AN	America (S)
Sango	SVO	Pr	NG		Africa
SENTANI	SOV	Po	GN	NA	N. Guinea
SHUSWAP	VSO	Pr	NG	AN	America (N)
SIROI	SOV	Po	GN	NA	N. Guinea
SULKA	SVO	Pr	GN	NA	N. Guinea
Tagalog		Pr	NG		Austronesian
Thai	SVO	Pr	NG	NA	Asia
Tiv	SVO	Pr	NG	NA	Africa
Tiwi	SVO	Pr	GN	AN	Australia
Tsou		Pr	NG		Austronesian
Turkish	SOV	Po	GN	AN	Asia
Vietnamese	SVO	Pr	NG	NA	Asia
WASKIA	SOV	Po	GN	NA	N. Guinea
Wati	SOV	Po	GN	NA	Australia
Welsh	VSO	Pr	NG	NA	Europe
Yapese	VSO	Pr	NG	NA	Austronesian
Yoruba	SVO	Pr	NG	NA	Africa
YUROK	SVO	Pr	GN	AN	America (N)
ZAPOTEC	VSO	Pr	NG	NA	America (C)
ZOQUE	SVO	Po	GN	AN	America (C)
Zulu	SVO	Pr	NG	NA	Africa

BIBLIOGRAPHY

A. GENERAL REFERENCES

Allwood, J., L.G. Anderson and O. Dahl (1977). Logic in Linguistics. Cambridge: University Press.

Anderson, J.M. and C. Jones (eds.) (1974). Historical Linguistics: Proceedings of the First International Conference on Historical Linguistics, Edinburgh, 2-7 September,1973. Two vols. Amsterdam: North Holland.

Bartsch, R. (1972). Adverbialsemantik. Habilitationsschrift, Berlin. Frankfurt-am-Main: Athenäumverlag.

Bartsch, R. (1973). "Semantics and Syntax of Number and Numerals". In: Kimball, J. (ed.) (1973).

Bartsch, R. (1976). The Grammar of Adverbials: a study in the semantics and syntax of adverbial constructions. Amsterdam: North-Holland.

Bartsch, R. and T. Vennemann (1972). Semantic Structures. A Study in the Relation between Syntax and Semantics. Frankfurt-am-Main: Athenäumverlag.

Bean, M. Callaway (1976). A study of the development of word-order patterns in Old English in relation to theories of word-order change. PhD dissertation, University of California, Los Angeles.

Bell, A. (1978). "Language Samples." In: Greenberg, J.H. (ed.) (1978).

Comrie, B. (1976). Aspect. Cambridge: University Press.

Comrie, B. (1981a) Languages of the Soviet Union. Cambridge: University Press.

Comrie, B. (1981b) Language Universals and Linguistic Typology. Syntax and Morphology. Oxford: Basil Blackwell.

Creider, C.A. (1979). "On the Explanation of Transformations." In: Givon, T. (1979).

Dixon, R.M.W. (1976). Grammatical Categories in Australian Languages. Linguistic Series No. 22. Australian Institute of Aboriginal Studies. Canberra: Humanities Press.

Dixon, R.M.W. (1977). "Where Have All the Adjectives Gone?" In: Studies in Language 1:1, pp. 19-80. Also in: Dixon, R.M.W. (1982).

Dixon, R.M.W. (1982). Where Have All the Adjectives Gone? and Other Essays in Semantics and Syntax. Berth, New York and Amsterdam: Mouton Publishers.

Dowty, D.R., R.E. Wall, and S. Peters (1981). Introduction to Montague Semantics. Dordrecht, Boston and London: David Reidel Publishing Company.

Everitt, B.S. (1977). The Analysis of Contingency Tables. London: Chapman and Hall.

Bibliography (continued)

Friederich, P. (1975). Proto-Indo-European Syntax: the Order of Meaningful Elements. Montana, Journal of Indo-European Studies.

Givón, T. (ed.) (1979). Syntax and Semantics. Volume 12: Discourse and Syntax. New York, San Francisco and London: Academic Press.

Greenberg, J.H. (1963). "Some Universals of Grammar with particular reference to the order of meaningful elements." In: Greenberg, J.H. (ed.) (1963).

Greenberg, J.H. (ed.) (1963). Universals of Language. Cambridge (Mass).and London: MIT Press.

Greenberg, J.H. (ed) (1978). Universals of Human Language. Volume 1: Method and Theory. Stanford (Calif): University Press.

Greenberg, J.H. (1980). "Circumfixes and Typological Change." In: Traugott, E.C. (ed.) (1980).

Hawkins, J.A. (1979). "Implicational Universals as Predictors of Word-order Change." In: Language 55, pp. 618-648.

Hawkins, J.A. (1980). "On Implicational and Distributional Universals of Word-order." In: Journal of Linguistics, volume 6, no. 2, pp. 193-235.

Heine, B. (1976). A Typology of African Languages Based on the Order of Meaningful Elements. Berlin: Dietrich Reimer Verlag.

Hoel, P.G. (1960). Elementary Statistics. New York and London: John Wiley and Sons, Inc.

Kamp, J.A.W. (1975). "Two Theories about Adjectives." In: Keenan, E.L. (ed.) (1975).

Keenan, E.L. (ed.) (1975). Formal Semantics of Natural Language. Cambridge: University Press.

Kimball, J. (ed.) (1973). Syntax and Semantics 2. New York, San Francisco and London: Academic Press.

Lehmann, W.P. (1973). "A Structural Principle of Language and its Implications." Language 49: 1, pp. 47-66.

Li, C.N. (ed.) (1975). Word-Order and Word-Order Change. Austin and London: University of Texas Press.

Montague, R. (1970a). "English as a Formal Language." In: Thomason, R.H. (ed.) 1970.

Montague, R.H. (1970b). "The Proper Treatment of Quantification in Ordinary English." In: Thomason, R.H. (ed.) (1970).

Prior, A.N. (1967). Past, Present and Future. Oxford: University Press.

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik (1972). A Grammar of Contemporary English. London: Longman.

Sapir, E. (1921). Language. An Introduction to the Study of Speech. New York: Harcourt, Brace and World.

Siegel, M. (1980). Capturing the Adjective. New York: Garland.

Bibliography (continued)

Silverstein, M. (1976) "Hierarchy of Features of Features and Ergativity." In: Dixon, R.M.W. (ed.) (1976).

Thomason, R.H. (ed.) (1970). Formal Philosophy. Selected Papers of Richard Montague. New Haven and London: Yale University Press.

Traugott, E.C., R. Labrum and S. Shepherd (ed.) (1980). Papers from the Fourth International Congress of Historical Linguistics. Amsterdam: John Benjamins B.V.

Vennemann, T. (1973). Explanation in Syntax. In: Kimball, J. (ed.) (1973).

Vennemann, T. (1974). "Topics, Subjects and Word-Order: From SXV to SVX via TVX." In: Anderson and Jones (1974).

Vennemann, T. (1975). "An Explanation of Drift." In: Li, C.N. (ed.) (1975).

Voegelin, C.F. and F.M. Voegelin (1977). Classification and Index of the World's Languages. New York, Oxford and Amsterdam: Elsevier.

B. LANGUAGE REFERENCES

Abrahams, R.C. (1933). The Grammar of Tiv. Kaduna: Government Printer.

Alphonse, E.S. (1956). Guaymí Grammar and Dictionary. Washington: U.S. Government Printing Office.

Armbruster, C.H. (1908). Initia Amharica. An Introduction to Spoken Amharic. Cambridge: University Press.

Arnott, D.W. (1970). The Nominal and Verbal Systems of Fula. Oxford: University Press.

Ashton, E.O. (1947). Swahili Grammar. 2nd ed. London: Longman.

Aspinion, R. (1953). Apprenons le berbere: initiation aux dialectes chleuhs. Rabat.

Bánhidi, B., Z. Jókay and D. Szabó (1965. Learn Hungarian. 2nd ed. Budapest: Tankönyvkiadó.

Biggs, B. (1973). Let's Learn Maori. A Guide to the Study of the Maori Language. 2nd ed. Wellington, Sydney and London: Reed Education, A.H. and A.W. Reed Ltd.

Boas, F. (1911). Handbook of American Languages. Part I. Washington: Government Printing Office.

Bowen, J.T. and T.J. Rhys-Jones (1960). Teach Yourself Welsh. London: English University Press Ltd.

Briggs, E. (1961). Mitla Zapotec Grammar. Mexico, D.F.: Instituto Lingüístico de Verano (SIL), and Centro de Investigaciones Antropológicas de México.

Canale, W.M. (1978). Word order change in Old English: base reanalysis in Generative Grammar. PhD dissertation: McGill University.

Capell, A. and H.E. Hinch (1970). Maung Grammar. The Hague and Paris: Mouton Ltd.

Chadwick, N. (1975). A Descriptive Study of the Djingili Language. Canberra: Australian Institute of Aboriginal Studies.

Chao, Yuen-ren (1968). A Grammar of Spoken Chinese. Berkeley and Los Angeles: University of California Press.

Chicago Linguistic Society (1973). You Take the High Node and I'll Take the Low Node. Papers from the Comparative Syntax Festival: The Differences between Main and Subordinate Clauses. Paravolume to Papers from the Ninth Regional Meeting.

Coulson, M. (1976). Teach Yourself Sanskrit. London: Hodder and Stoughton.

Cowan, H.K.J. (1965). Grammar of the Sentani Language. The Hague: Martinus Nijhoff.

Craig, C.G. (1977). The Structure of Jacaltec. Austin and London: University of Texas Press.

Crazzolara, J.P. (1960). A Study of the Logbara (Ma'di) Language. Oxford: University Press.

Day, C. (1973). The Jacaltec Language. Bloomington: Indiana University Publications.

DeFrancis, J. (1976). Beginning Chinese. 2nd ed. New Haven (Conn): Yale University Press.

Dirr, A. (1911). Rutul'skij Jazyk. Tbilisi: Tipografia Kanceljarii Namjestnika Jevo Imperatorskavo Veličestva na Kavkazje.

Dodds, R.W. (1977). Teach Yourself Malay. London: Hodder and Stoughton.

Douglas, W.H. (1959). An Introduction to the Western Desert Language of Australia. Oceanic Linguistic Monographs. Sydney: University of Sydney.

Drabbe, P. (1955). Spraakkunst van het Marind. Studia Instituti Anthropos. Vol. 11. Wien-Mödling: Drukkerij van het Missiehuis St. Gabriël.

Dunn, C.J. and S. Yannada (1958). Teach Yourself Japanese. London: Hodder and Stoughton.

Ergunvanlı, E.E. (1984). The Function of Word Order in Turkish Grammar. Berkeley, Los Angeles and London: University of California Press.

Ferrar, H. (1967). A French Reference Grammar. Oxford: University Press.

Gregores, E. and R. Súarez (1967). Description of Colloquial Guarani. The Hague and Paris: Mouton.

Grevisse, M. (1955). Le bon usage. Grammaire française avec des remarques sur la langue française d'aujourd'hui. 6me éd. Gembloux (Belgique): Editions J. Duculot, Paris: Libraire Orientaliste P. Geuthner.

Hacquard, F. and Dupuis, A. (1897). Manuel de la Langue Soñgay. Parlée de Timbouctou à Say dans la bouche du Niger. Paris: J. Maisonneuve.

Hagman, R.S. (1977). Nama Hottentot Grammar. Bloomington: Indiana University Publications.

Heath, J. (1978). Ngandi Grammar, Texts and Vocabulary. Canberra: Australian Institute of Aboriginal Studies, New Jersey: Humanities Press, Inc.

Hewitt, B.G. (1979). Abkhaz. Lingua Descriptive Series. Amsterdam: North Holland.

Hoff, B.J. (1968). The Carib Language. Phonology, Morphology, Texts and Word Index. The Hague: N.V. de nederlandsche book- en steendrukkerij v/h H.L. Smits.

Hoffmann, C. (1963). A Grammar of the Margi Language. Oxford: University Press.

Jensen, J.T. (1942). Yapese Reference Grammar. Honolulu: University Press of Hawaii.

Key, H. (1967). Morphology of Cayuvava. The Hague and Paris: Mouton.

Kimenyi, A. (1978). A Relational Grammar of Kinyarwanda. Berkeley, Los Angeles and London: University of California Press.

Bibliography (continued)

Kiss, K.E. (1981). "Structural Relations in Hungarian, a 'Free' Word Order Language." In: Linguistic Inquiry, Vol. 12, no. 2, pp.185-213.

Kraft. G.H. and A.H.M. Kirk-Greene (1973). Teach Yourself Hausa. London: Hodder and Stoughton.

Kroeber, A.L. and G.W. Grace (1960). The Sparkman Grammar of Luiseño. Berkeley and Los Angeles: University of California Press.

Kuipers, A.H. (1974). The Shuswap Language. The Hague and Paris: Mouton.

Kuno, S. (1972). "Functional Sentence Perspective: a case study from Japanese and English." Linguistic Inquiry 3: pp. 265-320.

Kuno, S. (1973). The Structure of the Japanese Language. Cambridge (Mass): MIT Press.

Kuno, S. (1978). "Japanese: a Characteristic OV Language." In: Lehman, W.P. (ed.) (1978).

Lazdina, T.B. (1966). Latvian. Teach Yourself Books. London: The English Universities Press Ltd.

Lehmann, W.P. (ed.) (1978). Syntactic Typology. Studies in the Phenomenology of Language. Hassocks, Sussex: The Harvester Press.

Lepsius, R. (1880). Nubische Grammatik. Berlin: Wilhelm Hertz.

Li, C.N. and S.A. Thompson (1973a). "Serial Verb Constructions in Mandarin Chinese." In: Chicago Linguistics Society (1973).

Li, C.N. and S.A. Thompson (1973b). "Historical Change of Word-Order: A Case Study of Chinese and its Implications." In: Anderson, J.M. and C. Jones (1974) (see Bibliography A).

Li, C.N. and S. Thompson (1975). "The Semantic Function of Word Order: a Case Study in Mandarin." In: Li, C.N. (ed.) (1975).

Li, C.N. and S. Thompson (1978). "An Exploration of Mandarin Chinese." In: Lehmann, W.P. (ed.) (1978).

Lorimer, D.L.R. (1935-8). The Burushaski Language. 3 vols. Oslo: Institutet for Sammenlignende Kulturforskning.

Lukas, J. (1937). A Study of the Kanuri Language. Grammar and Vocabulary. London, New York and Toronto: Oxford University Press, for the Institute of African Languages and Culture.

McCormack,W.(1966). Kannada. A Cultural Introduction to the Spoken Styles of the Language. With the assistance of M.G.Krishnamurthi. Madison,Milwaukee and London: The University of Wisconsin Press.

McGregor, D.E. and A.R.F. McGregor (1982). "Olo Language Materials." Pacific Linguistics. Series D - No. 42. Canberra: Australian National University.

McGregor, R.S.(1977). Outline of Hindi Grammar. Oxford: University Press.

McLendon, S. (1975). The Grammar of Eastern Pomo. Berkeley and Los Angeles: University of California Publications.

McLendon, S. (1978). "Ergativity in Eastern Pomo." International Journal of American Linguistics 44: pp. 1-9.

Mace, J. (1971). Modern Persian. Teach Yourself Books. London: Hodder and Stoughton.

Bibliography (continued)

Matteson, E. (1965). The Piro (Arawakan) Language. University of California Publications in Linguistics. Volume 45. Berkeley and Los Angeles: University of California Press.

Matthews, G.H. (1965). Hidatsa Syntax. The Hague: Mouton.

Metcalfe, C.D. (1975). Bardi Verb Morphology (North-western Australia). Canberra: Australian National University.

Milner, G.B. (1956). Fijian Grammar. Suva: Government Press.

Mitchell, T.F. (1956). An Introduction to Colloquial Egyptian Arabic. Oxford: University Press.

Müller, H. (1916). Erster Versuch einer Grammatik der Sulka-Sprache, Neu-Pommern (Sudsee). Anthropos. Vienna: Mechitaristen-Druckerei.

Murane, E. (1974). Daga Grammar. From Morpheme to Discourse. Norman: Summer School of Linguistics.

Needham, J.F. (1894). Outline of the Tai (Khamtî) Language. Rangoon: The Superintendent, Government Printing, Burma.

Noss, R.B. (1964). Thai Reference Grammar. Washington: Foreign Service Institute, Dept. of State.

Okell, J. (1969). A Reference Grammar of Colloquial Burmese. Oxford: University Press.

Osborne, C.R. (1974). The Tiwi Language. Canberra: Australian Institute of Aboriginal Studies.

Parker, G.J. (1969). Ayacucho Quechua Grammar and Dictionary. The Hague and Paris: Mouton.

Penzl, H. (1965). A Grammar of Pashto. A Descriptive Survey of the Dialect of Kandahar, Afghanistan. Washington: American Council of Learned Societies.

Praetorius, F. (1871). Grammatik der Tigriñasprache in Abessinien. Hauptsächlich in der Gegend von Aksum und Adoa. Halle: Verlag der Buchhandlung des Waisenhauses.

Quirk, R., S. Greenbaum, G. Leech and J. Svartvik (1972). A Grammar of Contemporary English. London: Longman.

Ramsay, G.G. (1897). Latin Prose Composition. Oxford: at the Clarendon Press.

Ramsey, M.M. (1956). A Textbook of Modern Spanish. Revised by R.K. Spaulding. New York: Holt, Rinehart and Winston.

Rausch, P.J. (1912). Die Sprachen von Südost-Bougainville, Deutsche Salomoninseln. Anthropos. Vienna: Mechitaristen-Buchdruckerei.

Reichle, V. (1981). Bawm Language and Lore. Berne, Frankfurt-am-Main and Las Vegas: Peter Lang.

Rivierre, J-C. (1980). La langue de Touho. Phonologie et grammaire du Cèmuhî (Nouvelle-Calédonie). Paris: Société d'Etudes Linguistiques et Anthropologiques de France (SELAF).

Robins, R.H. (1958). The Yurok Language. Grammar, Texts and Lexicon. University of California Publications in Linguistics. Volume 15. Berkeley and Los Angeles: University of California.

Bibliography (continued)

Ross, N. (1978). A Waskia Sketch and Vocabulary. With John Natill Paol. Canberra: Australian National University.

Rowlands, E.C. (1969). Yoruba. Teach Yourself Books. London: Hodder and Stoughton.

Ruffel Barlow, A. (1951). Studies in Kikuyu Grammar and Idiom. Edinburgh: Wm. Blackwood and Sons Ltd.

Ryecroft, D. and A.B. NGcobe (1981). Say it in Zulu. 2nd ed. London: School of Oriental and African Studies.

Samarin, W.J. (1967). A Grammar of Sango. The Hague: Mouton.

Schachter, P. and F.E. Otanes (1972). Tagalog Reference Grammar. Berkeley, Los Angeles and London: University of California Press.

Schneuker, C.L. (1962). Kâte Language Handbook. Madang: Lutheran Mission Press.

Scott, G. (1978). The Fore Language of Papua New Guinea. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.

Seiter, W.J. (1980). Studies in Niuean Syntax. New York and London: Garland Press.

Semeonoff, A. (1955). A New Russian Grammar. London and New York: Dent and Dutton.

Shipley, W.F. (1964). Maidu Grammar. Berkeley and Los Angeles: University of California Press.

Snyman, J.W. (1970). An Introduction to the !Xũ Language. Cape Town: A.A. Balkema.

Sofroniou, S.A. (1962). Modern Greek. Teach Yourself Books. London: Hodder and Stoughton.

Spears, R.A. (1965). The Structure of Faranah-Maninka. PhD thesis: Indiana University.

Srestacharya, I., J.N. Maskey and A. Hale (1971). Conversational Newari. Summer Institute of Linguistics.

Thompson, L. (1965). A Vietnamese Grammar. Seattle: University of Washington Press.

Tompa, J. (1968). Ungarische Grammatik. Budapest: Akademiai Kiadó.

Topping, D.M. (1973). Chamorro Reference Grammar. With the assistance of Bernardita C. Dungca. Honolulu: University Press of Hawaii.

Tucker, A.N. and J. Tompo Ole Mpaayei (1955). A Maasai Grammar with Vocabulary. London: Longmans, Green and Co.

Tung, T'ung-ho (1964). A Descriptive Study of the Tsou Language, Formosa. With the assistance of S.H. Wang, T.K. Khan, T.F. Cheung and Margaret M. Yan. Taipei (Taiwan): Institute of History and Philology, Academica Sinica. (Special Publications No. 48).

Underhill, R. (1976). Turkish Grammar. Cambridge (Mass): MIT Press.

Von Soden, W. (1952). Grundriss der Akkadischen Grammatik. Rome: Pontificium Institutum Biblicum.

Bibliography (continued)

Wells, M.A. (1978). Siroi Grammar. Canberra: Pacific Linguistics.

Welmers, W.E. (1968). Efik. Occasional Publication No. 11. Ibadan: Institute of African Studies, University of Ibadan.

Westermann, D. (1930). A Study of the Ewe Language. Oxford: University Press.

Whiteley, W.H. (1958). A Short Description of Item Categories in Iraqw. Kampala: East African Institute of Social Research.

Whitney, A.H. (1956). Finnish. Teach Yourself Books. London: English Universities Press Ltd.

Williams, W.L. (1923). First Lessons in Maori. New Zealand, Melbourne and London: Whitcombe and Tombs Ltd.

Williamson, K. (1965). A Grammar of the Kolokuma Dialect of Ijo. Cambridge: University Press.

Wonderley, W.M. (1951-2). Zoque: Phonology and Morphology. Reprinted from the International Journal of American Linguistics. Volume 17, Nos. 1, 2, 3 and 4, 1951, and volume 18, Nos. 1 and 4, 1952.

Additional Language References

Fortune, R.F. (1942) Arapesh New York: J.J. Augustin.

Grierson, G.A. (1911) A Manual of the Kashmiri Language 2 vols. Oxford: Unitersity Press.

James, D.L. (1972) Conversational Welsh Course(II) Llandybie(Caerffyrddin): Llyfrau'r Dryw.

Josephs, L.S. (1975) Palauan Reference Grammar Honolulu: University Press of Hawaii.

Lewis, G.L. (1953) Teach Yourself Turkish London: English Universities Press Ltd.

Lewis, G.L. (1967) Turkish Grammar Oxford: University Press.